IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPELLANT:	Alex Nugent	**EXAMINER:**	Joseph P. Hirl
SERIAL NO.:	10/735,934	**GROUP:**	2129
FILED:	12/15/2003	**ATTY DKT NO.:**	1000-1207

TITLE: **NANOTECHNOLOGY NEURAL NETWORK METHODS AND SYSTEMS**

Please forward all correspondence to:
ORTIZ & LOPEZ, PLLC
Patent Attorneys
P.O. Box 4484
Albuquerque, NM 87196-4484

Mail Stop Appeal Brief Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF FILED UNDER C.F.R. §1.192

Dear Sir:

This Appeal Brief is filed in support of the Notice of Appeal dated November 29, 2007 in the above-referenced application.

I. REAL PARTY IN INTEREST

Alex Nugent is the real parties in interest in the present invention, and is also the "Appellant" entitled to bring forward this appeal.

II. RELATED APPEALS AND INTERFERENCES

There are currently no related appeals and/or interferences related to the above-referenced patent application.

III. STATUS OF CLAIMS

This Appellant appeals the final rejection to claims 24-44 as set forth in the Final office action dated August 30, 2007. Claims 24-44 therefore constitute the appealed claims. The application was originally filed with 20 claims. In the first office action dated April 24, 2006, claims 1-20 were rejected. Appellant responded on May 23, 2006 to the first office action with an amendment in which original claims 1, 2, 14, 17, 20 were amended. In the second and final office action dated July 21, 2006, claims 1-20 were again rejected by the Examiner. The Appellant responded to this final office action on August 15, 2006 with an RCE (Request for Continued Examination) and amendments to claims 1, 5, 10, 14, 15, 16, 17, 18, 19, 20 and cancellation of claims 14, 18 and the presentation of newly submitted claims 21-23. The Examiner then responded with an office action dated October 25, 2006 in which claims 1-13, 15-17, and 19-23 were rejected. In response to the office action dated October, 26, 2006, the Appellant filed a response on December 18, 2006, in which 1, 10, 16, 17 and 20 were amended. The Examiner then issued a final office action dated March 2, 2007 again rejecting claims 1-13, 15-17, and 19-23 were again rejected. In response, to this second final office action, the Appellant filed another RCE on April 16, 2007 with supporting amendments in which claims 1-13, 15-17, and 19-23 were cancelled and new claims 24-44 submitted by amendment. The Examiner then issued an office action dated May 22, 2007 in which claims 24-44 were rejected. The Appellant responded by filing a response on

July 22, 2007 in which claims 24, 40 and 42 were amended. The Examiner issued a third final office on August 30, 2007 in response to the amendments filed on July 22, 2007. Appeal is now taken from the final office action dated August 30, 2007.

IV. STATUS OF AMENDMENTS

The amendments of claims 24, 40, 42 by Appellants prior to the Final office action dated August 30, 2007 are the claims that are now the subject of the appeal. Claims 24-44 remain pending in the appealed application.

The final rejection of claims 24-44 is the subject of this appeal.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The invention claimed in 24-44 is generally directed toward an electromechanical neural network system based on nanotechnology, which is taught by independent claims 24, 40 and 42. Dependent claims 25-39 (which depend from claim 24), and dependent claim 41 (which depends from claim 40) also teach various aspects and features of such an electromechanical neural network system based on nanotechnology. For purposes of summarizing the claimed subject matter, however, Appellant refers to independent claims 24, 40 and 42. The electromechanical neural network system is therefore shown and described with respect to FIGS. 1-38 and paragraphs [0001] to [0320] of Appellant's specification. Appellant's invention claims priority under 35 U.S.C. § 119(e) to provisional patent application Serial No. 60/488,860 filed July 18, 2003.

The language specifically distinguishing the independent claim 24 from the art of record is provided below:

24. An electromechanical neural network system based on nanotechnology, comprising:

an adaptive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution ;

a plurality of interconnected nanoconnections associated with said adaptive synaptic element, said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid die lectric solution; and

a voltage mechanism for applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors, thereby creating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

The claim limitations of Appellant's claim 24 are shown and described through Appellant's specification. For example, an adaptive synaptic element is shown in FIG. 2, which illustrates a schematic diagram illustrating a diode configuration as a neuron, in accordance with one embodiment. Such an adaptive synaptic element is described further in Appellant's paragraphs [0082] to [0086]. The claim limitations of a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution are illustrated, for example, in Appellant's FIG. 3, FIG. 37 and FIG. 38 and described throughout Appellant's specification. The claim limitations of a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode is also shown in Appellant's FIGS. 3, FIG. 37 and FIG. 38 and described in Appellant's specification. The claim limitations of wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent are described, for example, in Appellant's paragraphs [0099] – [00105]. The solvent and solution are additionally described, for example, in Appellant's paragraphs [00117] – [00122] of Appellant's specification with respect to Appellant's FIGS. 9-10. FIG. 9, for example, demonstrates the use of the solvent in forming the nanoconnections described in Appellant's specification and claims. The claim limitations of wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution is described, for example, in Appellant's paragraphs [0099] to [00100]. Examples of the claim limitations of a plurality of interconnected nanoconnections associated with said adaptive synaptic element are described in Appellant's

specification with respect to, for example, Appellant's FIGS. 2-3 and reference numerals 304, FIG. 5 and reference numerals 572, 574, 576, 578, and 580 and paragraphs [0093] – [0098]. The claim limitations of said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution are also described, for example, in Appellant's specification with respect to Appellant's FIG. 3, FIG. 37 and FIG. 38 and at various other locations throughout Appellant's specification. The claim limitations of said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution are also described in Appellant's specification with respect to Appellant's FIG. 3, FIG. 37 and FIG. 38 and at various other locations within Appellant's specification. Additionally, the claim limitations of a voltage mechanism for applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors are described with respect to Appellant's FIGS. 1-38 and the supporting sections thereof.

The claim limitations of a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap are discussed, for example, in Appellant's specification with respect to paragraphs [00272] - [00273].

The language specifically distinguishing the independent claim 40 from the art of record is provided below:

40. (Currently Amended) An electromechanical neural network system based on nanotechnology, comprising:

a resistive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said resistive synaptic element functions as an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode;

a plurality of interconnected nanoconnections associated with said resistive synaptic element, said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution;

a plurality of synapses associated with said resistive synaptic element, wherein said plurality of synapses comprises said plurality of interconnected nanoconnections of said resistive synaptic element and wherein each synapse among said plurality of synapses is independent of voltage polarization; and

a voltage mechanism for applying an AC electric field across said connection gap, whereby said AC electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said AC electric field across said connection gap so that said electromechanical neural network system adapts itself to the requirements of a given situation regardless of the initial state of said plurality of interconnected nanoconnections, wherein the longer the amount of time said AC electric field is applied across said connection gap and/or the greater the frequency or amplitude of said AC electric field applied across said connection gap, the more nanoconductors among said plurality of nanoconductors align and the stronger said interconnected nanoconnections among said plurality of nanoconnections become.

The claim limitations of a resistive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode are described at, for example, Appellant's paragraphs [0082] to [0086]. FIG. 37, for example, illustrates a pre-synaptic electrode 3702, a post-synaptic electrode 3704, and a network of nanoconnections 3708 formed between one or more respective input and output electrodes 3702 and 3704, in accordance with an alternative embodiment of the present invention. The claim limitations of wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent is described in Appellant's specification, for example, with respect to paragraphs [0099] – [00100] and Appellant's FIG. 3, FIG. 37 and FIG. 38 and at various other locations within the specification. The claim limitations of wherein said resistive synaptic element functions as an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode are indicated by, for example FIG. 10 and paragraph [00125] which indicates that "…as indicated at block 1010, as the electric field is applied across the connection gap, the more the

nonconductor(s) will align and the stronger the connection becomes...Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block 1012...As illustrated at block 1014, the resistance of the connection can be maintained or lowered by selective activations of the connections...In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain in response to Long Term Depression, or LTD." Examples of pre- and post-synaptic electrodes are shown in FIG. 37-38 of Appellant's specification.

The claim limitations of a plurality of synapses associated with said resistive synaptic element, wherein said plurality of synapses comprises said plurality of interconnected nanoconnections of said resistive synaptic element and wherein each synapse among said plurality of synapses is independent of voltage polarization is described in Appellant's specification with respect to, for example, FIG. 27-38, and the supporting specification thereof.

Additionally, the claim limitations of a voltage mechanism for applying an AC electric field across said connection gap, whereby said AC electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said AC electric field across said connection gap so that said electromechanical neural network system adapts itself to the requirements of a given situation regardless of the initial state of said plurality of interconnected nanoconnections, wherein the longer the amount of time said AC electric field is applied across said connection gap and/or the greater the frequency or amplitude of said AC electric field applied across said connection gap, the more nanoconductors among said plurality of nanoconductors align and the stronger said interconnected nanoconnections among said plurality of nanoconnections become, are described, for example, in Appellant's specification with respect to FIGS. 18-34 and their related description. Numerous examples of voltage mechanisms are

illustrated in described in Appellant's specification. For example, FIG. 15 shows various electrically conducting input and output lines in a multi-layer structure. Voltage across these lines constitutes a voltage mechanism. FIG. 14 shows a similar but different structure. A voltage mechanism is also shown in FIG. 13 with respect to +/- inputs and outputs and so forth. In FIG. 12, any one of the components 1213, 1214, 1215, 1216, 1222 etc constitute voltage mechanisms.

The language specifically distinguishing the independent claim 42 from the art of record is provided below:

42. (Currently Amended) A method of forming an electromechanical neural network system based on nanotechnology, comprising:

providing a liquid dielectric solution comprising a mixture of a dielectric solvent and a plurality of nanoconductors, wherein each nanoconductor among said plurality of nanoconductors is suspended and free to move about in said liquid dielectric solution;

forming a connection gap between at least one pre-synaptic electrode and said at lease post-synaptic electrode;

configuring a connection network to comprise said plurality of nanoconductors suspended and free to move about in said liquid dielectric solution located within said connection gap formed between said at least one pre-synaptic electrode and said at least one post-synaptic electrode, wherein said connection network comprises an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode;

configuring said connection network to comprises a plurality of interconnected nanoconnections with said connection network, wherein said plurality of interconnected nanoconnections comprise said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution;

providing a plurality of synapses from said connection network, wherein said plurality of synapses comprises said plurality of interconnected nanoconnections of said connection network and wherein each synapse among said plurality of synapses is independent of voltage polarization; and

applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

The claim limitations of providing a liquid dielectric solution comprising a mixture of a dielectric solvent and a plurality of nanoconductors, wherein each nanoconductor among said plurality of nanoconductors is suspended and free to move about in said liquid dielectric solution is disclosed, for example, in Appellant's

FIG. 9 with respect to blocks 902, 904, 906, 908 and 910 and at various other locations within Appellant's specification. The claim limitations of forming a connection gap between at least one pre-synaptic electrode and said at lease post-synaptic electrode are shown in a number of locations within Appellant's specification, such as, for example, Appellant's FIG. 3 with respect to paragraphs [0086] – [0093]. Additionally, FIGS. 37-38 demonstrate aspects of a connection gap between at least one pre-synaptic electrode and said at lease post-synaptic electrode.

The claim limitations of configuring a connection network to comprise said plurality of nanoconductors suspended and free to move about in said liquid dielectric solution located within said connection gap formed between said at least one pre-synaptic electrode and said at least one post-synaptic electrode, wherein said connection network comprises an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode are also shown and described in a number of locations within Appellant's specification. For example, the connection network is shown in FIG. 3 and described with respect to paragraphs [0086] – [0093]. The connection gap is also shown and described in FIG. 3 with respect to paragraphs [0086] – [0093]. FIGS. 37-38 demonstrate aspects of a connection gap between at least one pre-synaptic electrode and said at lease post-synaptic electrode.

The claim limitations of configuring said connection network to comprises a plurality of interconnected nanoconnections with said connection network, wherein said plurality of interconnected nanoconnections comprise said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution are also shown and described in numerous locations within Appellant's specification such as with respect to FIG. 3 with respect to paragraphs [0086] – [0093] and FIGS. 37-38 along with FIGS. 9-10.

The claim limitations of providing a plurality of synapses from said connection network, wherein said plurality of synapses comprises said plurality of

interconnected nanoconnections of said connection network and wherein each synapse among said plurality of synapses is independent of voltage polarization are also demonstrated in a number of locations within Appellant's specification such as, for example, FIG. 9-10 and FIG. 11 and FIGS. 4, 5, 7, and 8 with respect to the description of FIG. 9-10 and FIG. 11 and FIGS. 4, 5, 7, and 8 found in Appellant's specification. The claim aspect of wherein said plurality of synapses is independent of voltage polarization is described, for example, in FIGS. 27-38 and the supporting description thereof.

The claim limitations of applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap are shown for example, with respect to FIGS. 9-10 and paragraphs [00117] – [00133]. The use of a dielectric solution in association with the electric field and the plurality of nanoconductors is dielectrophoresis. An example of a dielectrophoretic force is discussed, for example, in Appellant's paragraphs [00271] – [00272].

For a full understanding and appreciation of the concepts behind Appellant's claims, it is suggested that one read the entire specification in its entirety.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Issues #1, #2, #3, #4, #5, #6, #7, #8, #9, #10 and #11 below constitute the grounds of rejection to be reviewed on appeal. A concise statement of each ground of rejection presented for review is listed below with respect to Issues #1, #2, #3, #4, #5, #6, #7, #8, #9, #10 and #11.

-- ISSUE #1: Whether claims 24-44 comply with the enablement requirement of 35 U.S.C. 112, first paragraph.

-- ISSUE #2: Whether claims 24-44 comply with the written description requirement of 35 U.S.C. 112, first paragraph.

-- ISSUE #3: Whether Appellant's concept of a liquid dielectric solution comprising a mixture of a plurality of nanoconductors and a liquid dielectric solvent wherein a plurality of nanoconductors are free to move about in a dielectric solution and such a solution is disposed between two electrodes is anticipated by Paul M. Adriani and Alice P. Gast in the article entitled "Electric-field-induced aggregation in dilute colloidal suspensions" published in 1990 by the Faraday Discussions of the Chemical Society, hereinafter referred to as the Adriani reference.

-- ISSUE #4: Whether Appellant had defined the term "nanotechnology" in Appellant's specification.

-- ISSUE #5: Whether Appellant had defined the term "solvent" in Appellant's specification.

-- ISSUE #6: Whether Appellant had defined the term "liquid dielectric solution" in Appellant's specification.

-- ISSUE #7: Whether the prior art cited by the Examiner conclusively establishes that the invention of the Appellant will not function as a neural network.

-- ISSUE #8: Whether claims 24-44 lack patentable utility under 35 U.S.C. § 101.

-- ISSUE #9: Whether the Appellant has in fact disclosed the practical application for the invention and whether the rejection to claims

24-44 under 35 U.S.C. § 112 should be withdrawn because the Appellant has NOT failed as a matter of law to satisfy 35 U.S.C. § 101.

-- ISSUE #10: Whether claims 24-44 fail to identify an invention (neural network) that can be evaluated under the conditions of novelty or nonobviousness with respect to 35 U.S.C. § 102 or § 103.

-- ISSUE #11: Whether the requirement for information under 35 U.S.C. §112, 37 C.F.R. § 1.105 was proper.

VII. ARGUMENT

APPLICABLE LEGAL STANDARDS

35 U.S.C. §112, 37 C.F.R. § 1.105 Requirements for information

One of the relevant statutes cited in rejecting Appellant's claims is 37 C.F.R. § 1.105, Requirements for information, which indicates the following:

(a)

(1) In the course of examining or treating a matter in a pending or abandoned application filed under 35 U.S.C. 111 or 371 (including a reissue application), in a patent, or in a reexamination proceeding, the examiner or other Office employee may require the submission, from individuals identified under § 1.56(c), or any assignee, of such information as may be reasonably necessary to properly examine or treat the matter, for example:

(i) *Commercial databases*: The existence of any particularly relevant commercial database known to any of the inventors that could be searched for a particular aspect of the invention.

(ii) *Search*: Whether a search of the prior art was made, and if so, what was searched.

(iii) *Related information*: A copy of any non-patent literature, published application, or patent (U.S. or foreign), by any of the inventors, that relates to the claimed invention.

(iv) *Information used to draft application*: A copy of any non-patent literature, published application, or patent (U.S. or foreign) that was used to draft the application.

(v) *Information used in invention process*: A copy of any non-patent literature, published application, or patent (U.S. or foreign) that was used in the invention process, such as by designing around or providing a solution to accomplish an invention result.

(vi) *Improvements*: Where the claimed invention is an improvement, identification of what is being improved.

(vii) *In Use:* Identification of any use of the claimed invention known to any of the inventors at the time the application was filed notwithstanding the date of the use.

(viii) *Technical information known to Appellant*. Technical information known to Appellant concerning the related art, the disclosure, the claimed subject matter, other factual information pertinent to patentability, or concerning the accuracy of the examiner's stated interpretation of such items.

(2) Where an assignee has asserted its right to prosecute pursuant to § 3.71(a) of this chapter, matters such as paragraphs (a)(1)(i), (iii), and (vii) of this section may also be applied to such assignee.

(3) Requirements for factual information known to Appellant may be presented in any appropriate manner, for example:

(i) A requirement for factual information;

(ii) Interrogatories in the form of specific questions seeking Appellant's factual knowledge; or

(iii) Stipulations as to facts with which the Appellant may agree or disagree.

(4) Any reply to a requirement for information pursuant to this section that states either that the information required to be submitted is unknown to or is not readily available to the party or parties from which it was requested may be acc epted as a complete reply.

(b) The requirement for information of paragraph (a)(1) of this section may be included in an Office action, or sent separately.

(c) A reply, or a failure to reply, to a requirement for information under this section will be governed by §§ 1.135 and 1.136.

35 U.S.C. §112, First paragraph (Enablement and Written Description Requirements)

One of the relevant statutes cited in rejecting Appellant's claims is 35 U.S.C. §112, which indicates that the specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor of carrying out his invention.

35 U.S.C. §101

One of the relevant statutes cited in rejecting Appellant's claims is 35 U.S.C. §101, which indicates that whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title.

Constructive Reduction to Practice

A legal concept that must be considered in the context of the rejection by the Examiner to Appellant's claims is the concept of a "constructive reduction to practice". MPEP 2138.05 "Reduction to Practice" [R-5] - 2100 Patentability, indicates that a reduction to practice may be an actual reduction or a <u>constructive</u>

reduction to practice which occurs when a patent application on the claimed invention is filed. The filing of a patent application serves as conception and constructive reduction to practice of the subject matter described in the application. Thus, the inventor need not provide evidence of either conception or actual reduction to practice when relying on the content of the patent application. *Hyatt v. Boone*, 146 F.3d 1348, 1352, 47 USPQ2d 1128, 1130 (Fed. Cir. 1998). Additionally, constructive reduction to practice: "[O]ccurs upon the filing of a patent application on the claimed invention." *Brunswick Corp. v. U.S.*, 34 Fed. Cl. 532, 584 (1995).

35 U.S.C. §102(b)

Another relevant statute cited in rejecting Appellants' claims is 35 U.S.C. §102(b), Conditions for patentability; novelty and loss of right to patent. Section (b) is the basis of the rejections rendered by the examiner. Under 35 U.S.C. §102, section (b), a person is be entitled to a patent unless:

> (b) - the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States ---

Prima Facie Anticipation

The Commissioner of Patents and Trademarks, acting through examining officials, bears the initial duty of supplying the factual basis supporting a rejection of a patent application, including a rejection based on anticipation. *In re Warner*, 379 F.2d 1011, 154 USPQ 173, 178 (C.C.P.A. 1967), *cert. denied*, 389 U.S. 1057 (1968). The courts have interpreted this initial duty as placing on the Commissioner and the examiner the burden of presenting a *prima facie* case of anticipation. *See In re King*, 801 F.2d 1324, 1327, 231 USPQ 136, 138-39 (Fed. Cir. 1986); *In re Wilder*, 429 F.2d 447, 450, 166 USPQ 545, 548 (C.C.P.A. 1970). As stated by the Board in *In re Skinner*, 2 USPQ 2d 1788, 1788-9 (B.P.A.I. 1986), "[i]t is by now well settled that the burden of establishing a *prima facie* case of anticipation resides with the Patent and Trademark Office."

A general definition of *prima facie* unpatentability is provided at 37 C.F.R. §1.56(b)(2)(ii):

> A *prima facie* case of unpatentabilit y is established when the information *compels a conclusion* that a claim is unpatentable under the preponderance of evidence, burden-of-proof standard, giving each term in the claim its broadest reasonable construction <u>consistent with the specification</u>, and before any consideration is given to evidence which may be submitted in an attempt to establish a contrary conclusion of patentability. (Emphasis added.)

"Anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration." *W.L. Gore & Associates v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303, 313 (Fed. Cir. 1983) (citing *Soundscriber Corp. v. United States*, 360 F.2d 954, 960, 148 USPQ 298, 301 (Ct. Cl.), *adopted*, 149 USPQ 640 (Ct. Cl. 1966)), *cert. denied*, 469 U.S. 851 (1984). Thus, to anticipate the Appellants' claims, the reference cited by the Examiner must disclose each element of the respective claims that they are being recited for. "There must be no difference between the claimed invention and the reference disclosure, as viewed by a person of ordinary skill in the field of the invention." *Scripps Clinic & Research Foundation v. Genentech, Inc.*, 927 F.2d 1565, 18 USPQ 2d 1001, 1010 (Fed. Cir. 1991).

To overcome the anticipation rejection, the Appellants need only demonstrate that not all elements of a *prima facie* case of anticipation have been met, *i. e.*, show that the reference cited by the Examiner does not disclose every element in each of the Appellants' claims associated with the relevant reference used for their rejection. "If the examination at the initial state does not produce a prima face case of unpatentability, then without more the Appellant is entitled to grant of the patent." *In re Oetiker*, 977 F.2d 1443, 24 USPQ 2d 1443, 1444 (Fed. Cir. 1992).

35 U.S.C. §103(a)

The relevant statute cited in rejecting Appellants' claims is 35 U.S.C. §103(a),

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Requirements for Prima Facie Obviousness

The obligation of the examiner to go forward and produce reasoning and evidence in support of obviousness is clearly defined at M.P.E.P. §2142:

> "The examiner bears the initial burden of factually supporting any *prima facie* conclusion of obviousness. If the examiner does not produce a *prima facie* case, the applicant is under no obligation to submit evidence of nonobviousness."

The U.S. Supreme Court ruling of April 30, 2007 (*KSR Int'l v. Teleflex Inc.*) states:

"The TSM test captures a helpful insight: A patent composed of several elements is not proved obvious merely by demonstrating that each element was, independently, known in the prior art. Although common sense directs caution as to a patent application claiming as innovation the combination of two known devices according to their established functions, it can be important to identify a reason that would have prompted a person of ordinary skill in the art to combine the elements as the new invention does."

"To facilitate review, this analysis should be made explicit."

The U.S. Supreme Court ruling states that it is important to identify a *reason* that would have prompted a person to combine the elements and to make that analysis *explicit*. MPEP §2143 sets out the further basic criteria to establish a *prima facie* case of obviousness:

1. a reasonable expectation of success; and

2. the teaching or suggestion of <u>all</u> the claim limitations by the prior art reference (or references when combined).

It follows that in the absence of such a *prima facie* showing of obviousness by the Examiner (assuming there are no objections or other grounds for rejection) and of a *prima facie* showing by the Examiner of a *reason* to combine the references, an applicant is entitled to grant of a patent. Thus, in order to support an obviousness rejection, the Examiner is obliged to produce evidence compelling a conclusion that the basic criterion has been met.

APPELLANT'S ARGUMENTS REGARDING ISSUE #1 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Claims 24-42 comply with the enablement requirement of 35 U.S.C. 112, first paragraph.

Claims 24-44 were rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The Examiner argued that the claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

The Examiner argued that proof of a constructive reduction to practice requires sufficient disclosure under the "how to use" and "how to make" requirements of 35 U.S.C. 112, first paragraph. In support of this argument, the Examiner cited Kawai v. Metlesics, 480 F.2d 880, 886, 178 USPQ 158, 163 (CCPA 1973). The Examiner indicated that any analysis of whether a particular claim is supported by the disclosure in an application requires a determination of whether that disclosure, when filed, contained sufficient information regarding the subject matter of the claims as to enable one skilled in the pertinent art to make and use the claimed invention. The Examiner also indicated that the standard for determining whether the specification meets the enablement requirement was cast in the Supreme Court decision of Mineral Separation v. Hyde, 242 U.S. 261, 270 (1916) which postured the question: is the experimentation needed to practice the invention undue or unreasonable? The Examiner indicated that standard is still the one to be applied. The Examiner cited In re Wands, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir.1988) in making this argument. The Examiner asserted that even though the statute does not use the term "undue experimentation," it has been interpreted to require that the claim invention be enabled so that any person skilled in the art can make and use the invention without undue experimentation. In re Wands, 858 F .2d at 737, 8 USPQ2d at 1404 (Fed. Cir. 1988) ("The test of enablement is whether one reasonably skilled in the art could make or use the invention from the disclosures in the patent coupled with information known in the art without undue experimentation.").

From In re Wands, the Examiner argued that the factors to be considered in determining whether a disclosure would require undue experimentation are considered as follows:

[The Examiner noted, however, that in responding, the Examiner uses the art provided by the Appellant in response to the office action dated July 23, 2007; Hong et al., Controllable Capture of Au Nano-Particles by Using Dielectrophoresis, December 2004, referred to as Hong. The Examiner admitted that such art is not prior art, but asserted that it provides an insight to the level of skill existing at the time of the Appellant's filing dated December 15, 2003.

The Examiner indicated that the subject art by Hong was prepared by S. H. Hong, H. K. Kim, B. C. Kim, Y. S. Choi of the Department of Electronics and Computer Engineering, Korea University; J. S. Hwang and D. Ahn of the Institute of Quantum Information Processing and Systems, University of Seoul, S. K. Kwak and D. J. Ahn of the Department of Chemical and Biochemical Engineering, Korea University; and S. W. Hwang of the Department of Electronics and Computer Engineering, Korea University and such art was published in the Journal of the Korean Physics Society. The Examiner argued that such a group of individuals, referred to as University Group, cannot be considered at the level of one of ordinary skill in the art.]

A. The quantity of experimentation necessary: The Examiner argued that the University Group, in the art (i.e., the Hong reference) printed one year after the filing of the Appellant's application write of the difficulty in controllably capturing of Au nano-particles by using dielectrophoresis (the Examiner cited page s665, column 2, lines 3-24 of the Hong reference). The Examiner asserted that the Appellant in the specification fails to identify the "how to make" electromechanical neural network that is expected to come about from a plurality of nanoconductors in a dielectric solution such that one of ordinary skill in the art can replicate the invention ... one of ordinary skill in the art is not at the level of the University Group and that University Group had difficulty achieving their result at their level of skill one year after the Appellant filed the application. The Examiner asserted that the quantity of experimentation by one of ordinary skill in the art is undue, etc.

The Appellant respectfully disagrees with this assessment. First, the Hong reference does not constitute "prior art" for purposes of any of the legal standards such as 35 U.S.C. 112, 101, 112, 102 or 103 discussed herein. Thus, the use of the Hong reference by the Examiner to argue any level skill in the art at the time of filing of the Appellant's specification is irrelevant. Second, the Hong reference was provided originally to demonstrate NOT an electromechanical neural network, but merely to point out that aspects of the use of dielectrophoresis to trap nanoconductors in a dielectric solution have been proven out. In this regard, the Hong reference was cited as a basis for demonstrating that aspects of dielectrophoresis and nanoconductors and electrodes as utilized by Appellant's invention have been demonstrated, thereby satisfying the Examiner's request for information and test data, even though Appellant submits that such information is not necessary, and such a request by the Examiner is improper, given that Appellant's specification already constitutes proof of a constructive reduction to practice requires sufficient disclosure under the "how to use" and "how to make" requirements of 35 U.S.C. 112, first paragraph.

Appellant's disclosure is more than sufficient for purposes of 35 U.S.C 112 for several aspects. First, Appellant's invention is actually a quite simple concept, while at the same time, is unique and unobvious over the prior art. Appellant submits that "undue experimentation" is not required for one of ordinary skill in the art to replicate Appellant's invention. Simply following the methodology disclosed in Appellant's specification and the proper machinery is sufficient to replicate Appellant's invention. Appellant's specification provides numerous examples of "how to use" and "how to make" Appellant's claimed invention. For example, FIGS. 9-10 of Appellant's specification illustrate actual method steps that one skilled in the art can readily follow to implement Appellant's invention.

Appellant notes that the trapping of nanoconductors in a connection is not the issue here as this is a currently well-developed theory complete with multiple textbooks on the subject. Thus, the quantity of experimentation necessary to replicate this aspect of Appellant's invention is as simple as buying a textbook on dielectrophoresis. All of the unique aspects of Appellant's invention (e.g., microelectronics, patterning of electrode gaps, dielectrophoresis, neural networks,

etc) are well-developed arts that have been implemented by many groups across the globe. It is the unique combination of these elements that renders Appellant's invention new and unobvious over this art. It is therefore incorrect for the Examiner to assert that these components could not be combined to form Appellant's electromechanical neural network. Thus, the quantity of experimentation necessary to replicate Appellant's invention is not difficult, given that each of these areas are well-established. The Hong reference merely demonstrates an example of one aspect of Appellant's invention (i.e., trapping of nanoconductors), and does not represent a complete survey of the field of dielectrophoresis.

Appellant reiterates that the University Group and hence, the Hong reference, was not an attempt to "replicate" Appellant's invention, but merely provides evidence that certain aspects of Appellant's invention (but not Appellant's actual invention) have been demonstrated. Because these aspects have been demonstrated, then it is reasonable to conclude that Appellant's invention is also implementable. Regardless of any difficulty in controllably the capturing of Au nano-particles by using dielectrophoresis by the University Group (the Examiner cited page s665, column 2, lines 3-24 of the Hong reference) as argued by the Examiner, the Hong reference does provide evidence that nanoconnections such as Appellant's nanoconnections can be grown directly within a connection gap rather than at electrodes as the Examiner had argued in prior office actions.

Regarding the difficulty in controlling the capturing of Au nano-particular by using dielectrophoresis by the University Group as pointed out by the Examiner, the Appellant notes that the only thing that is difficult as stated by the University Group (see second sentence of Hong abstract) is to find the actual time point of capture by measuring the current across the gap. This is because in the specific instance of the Hong combination of nanoparticles (i.e., gold), the electrode material (i.e., also gold) and the solution of Hong, in addition to a self-assembled mono-layer of 1,8 octanedithiol molecules (see page S666, second sentence, section II of Hong), a trapped nanoparticle is prevented from conducting electricity WHILE IN SOLUTION. This is again is a particular instance of the specification combination of materials used in the Hong reference, but is not evidence of the quantity of experimentation

necessary to replicate Appellant's invention, because there are abundant examples of other nanoparticle/electrode combination solutions. For example, see the following reference, which is included herewith in the IX. EVIDENCE APPENDIX:

N. Mureau, E. Mendoza, K. Hoettges, S.R.P. Silva and M.P. Hughes, "In-situ and real time determination of metallic and semiconducting SWNT in suspension via dielectrophoresis", Appl. Phys. Lett., 88, (2006) 243109.

In this specific example, known as the Mureau reference, the problems noted by the University Group with respect do not exist and actually measuring current across the gap while in solution is specifically used by Mureau to ascertain the ratio of conducting to non-conducting nanotubes. Appellant's note that the Mureau reference is not a prior art reference with respect to Appellant's invention as the Mureau reference post-dates the date of filing of Appellant's patent application.

In order to demonstrate the of Appellant's invention involving nonconductors in a solution, and ultimately, Appellant's overall nanometer scale based device, the Appellant had previously requested the Examiner to consider the case of a charged nanoparticle suspended in a liquid dielectric solution between two electrodes. Appellant provided the Examiner with the following discussion to demonstrate that the nanoconnections formed in a solution and attracted to a connection gap (rather than electrodes) is based on practical principals and can function as a neural network. (Note: Appellant is, of course, not asserting that his nanoparticles must be charged, but just providing a helpful illustration).

The Appellant previously requested that the Examiner consider the situation wherein an <u>alternating</u> electric field is applied across the electrodes. Because the field is alternating, the charged particle is equally attracted and repelled to/from both electrodes. Nanoparticles are moved by a dipole induced force, e.g., the dielectrophoretic force. Such a force can be described mathematically as follows:

$$\vec{F}_{dep} = 2\pi r^3 \varepsilon_0 \varepsilon_m \mathrm{Re}\left[\frac{\varepsilon_p^* - \varepsilon_m^*}{\varepsilon_p^* + 2\varepsilon_m^*}\right]\nabla E^2$$

The force is dependent on the gradient of the square of the magnitude of the electric field. It is standard physics knowledge that the electric field inside a conductor is zero. As soon as a conducting nanoparticle touches an electrode there can be no electric field between the nanoparticle and the electrode because their electric potentials are equal. Thus, the moment the particle actually touches the electrode is the moment the dielectrophoretic force is in essence turned off. One can then argue that the particle was in fact attracted to the electrode gap and not the electrode.

In order to demonstrate these concepts, the Appellant previously referred to the Hong reference and the following figure from the Hong reference for the convenience of the Examiner:

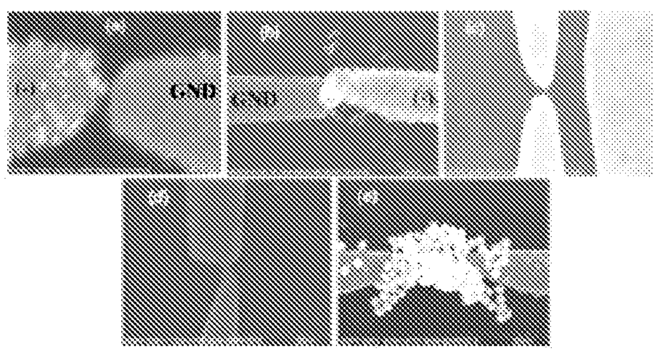


Fig. 2. (a) SEM image of captured Au nano-particles of diameter 20 nm under DC bias of - 2.05 V. (b) SEM image of captured Au nano-particles of diameter 40 nm under DC bias of - 3.3 V. (c) SEM image of captured Au nano-particles of diameter of 20 nm under AC bias of 1 V (peak-to-peak), 100 kHz, for 8 seconds. (d) SEM image of captured Au nano-particles of diameter 30 nm under AC bias of 4 V, 1 MHz for 40 seconds. (e) SEM image of captured Au nano-particles of diameter 50 nm under AC bias of 5 V, 1 MHz, for 300 seconds.

The Hong et al reference, which post dates Appellant's application and thus is not a proper prior art reference, merely demonstrates that particles are attracted only to the electrode gap. One needs only compare Fig. 2(b) to Fig. 2(d) of the Hong et al reference. In Fig. 2(b) a non-alternating (i.e., static) voltage is applied across electrode terminals. Because the gold nanoparticles obtain a net positive charge when placed in solution, they are attracted only to the negative electrode.

Contrast this with the case shown by Fig. 2(d) of the Hong et al reference where the applied voltage is now alternating. Note how the particles are only attracted to the electrode gap and do not accumulate on either electrode. Fig. 2(e) shows what happens when the magnitude of the alternating voltage is increased by a volt and allowed to continue for 260 seconds longer. Note that the accumulation of particles is directed to the electrode gap. However, the abundance of particles requires that they necessarily touch the electrodes. It is very clear by these pictures that particles are attracted to the gap and not the electrodes themselves.

The Appellant had previously encouraged the Examiner to explore the field of negative dielectrophoresis whereby certain electrode geometries can be used to trap non-charged particles while never touching the electrodes. The Appellant also previously encouraged the Examiner to explore the area of laser tweezers because both of these fields directly manipulate charge neutral particles without ever touching them with electrodes. The Examiner's final office action dated August 30, 2007 did not indicate that the Examiner had explored these areas.

In any event, the comments by the University Group and hence, the Hong reference, does not provide evidence that one skilled in the art would require "undue experimentation" to replicate Appellant's invention. Again, the Hong reference was merely provided as an attempt to satisfy the Examiner's unnecessary request for information under 37 CFR 1.105, in that the Hong reference shows that nanoconductors can be attracted to a connection gap rather (rather than electrodes). Such a connection gap is clearly disclosed in Appellant's specification in a number of places, such as, for example, the connection gap of FIG. 3, FIGS. 9-10, and in FIGS. 37-38 and at numerous other locations within Appellant's specification.

The Appellant therefore submits that that Appellant's specification does in fact identify "how to make" and "how to use" Appellant's electromechanical neural network and that "undue experimentation" by those skilled in the art is not necessary to replicate Appellant's invention.

B. The amount of direction and guidance: The Examiner argued that the rationale of item A above applies. The Examiner argued that the Appellant in the

Remarks, dated July 23, 2007, page 10, lines 10-13 stated: "Appellant presents the following discussion to demonstrate that the nanoconnections formed in a solution and attracted to a connection gap (rather than electrodes) is based on practical principles and can function as a neural network." The Examiner argued that any one of the pictures shown in Fig. 2 of Hong do not resemble the proposed network of Appellant identified in Fig. 3 of the specification. The Examiner asserted that Appellant's direction in the "how to make" such that one of ordinary skill in the art can replicate the invention has not been found in the specification.

The Appellant respectfully disagrees with this assessment. The Examiner seems to have confused the reasoning for why the Hong reference was provided to the Examiner in the first place. Again, as argued above with respect to A, the Hong reference was not provided to the Examiner to demonstrate the pictures shown in FIG. 2 of Hong resemble Appellant's neural network identified in FIG. 3 of Appellant's specification. The Hong reference does not teach a neural network and thus does not offer any guidance regarding neural networks and neural network principals. The Hong reference is irrelevant when it comes to neural networks. In this regard, it is uncertain why the Examiner is even arguing 35 U.S.C. 112 requirements based on the Hong reference, when the Hong reference does not even provide for a teaching of neural networks, which are key components of Appellant's invention.

The Hong reference as indicated above in the response to item A above was provided to the Examiner in response to the Examiner's request for more information merely to demonstrate certain key aspects of Appellant's invention such as, for example, the attraction of nanoconductors to the connection gap and not the electrodes. By demonstrating that Appellant's nanoconductors are attracted to the connection gap and not the electrodes, this demonstrates an important aspect of Appellant's electromechanical neural network, and thus is evidence of how Appellant's neural network can be configured. Even without the Hong reference, however, Appellant's specification provides numerous examples of direction and guidance and hence, "how to make" Appellant's invention, and additionally how one of ordinary skill in the art can replicate the invention (e.g., such as FIGS. 9-10 of Appellant's specification). There is thus not a need to provide the Examiner with

evidence of "test data" because Appellant's invention properly demonstrates the claimed invention.

Appellant again notes that all of the unique aspects of Appellant's invention (e.g., microelectronics, patterning of electrode gaps, dielectrophoresis, neural networks, etc) are well-developed arts that have been implemented by many groups across the globe. It is the unique combination of these elements that renders Appellant's invention new and unobvious over this art. It is therefore unreasonable to argued that the there is not a sufficient amount of direction and guidance in the art. The Examiner's argument seems to be analogous to arguing that there is not sufficient direction and guidance in the art to build a car given that a transmission exists, engines exist, tires exist, etc. The Hong reference is merely an example of one piece of the art. Hong does not teach neural networks. It teaches that nanoparticles can be attracted to an electrode gap. Appellant does not understand why the Examiner is pointing to Hong as an example of "the amount of direction and guidance" in the art when Hong is not even prior art with respect to Appellant's invention.

C. The presence or absence of working examples: The Examiner indicated that in the Office Action dated May 22, 2007, a request was made for test data and pictures of test documentation related to the invention (the Examiner referred to ¶ 4. above) and that none has been provided. Further, the Examiner asserted that no statement by the Appellant has been made that such material has not been provided because it does not exist ... an attempt to substitute using the art of Hong was made. The Examiner argued that based on the action by the Appellant, it can only be assumed that working examples do not exist ... the Examiner asserted no enablement.

The Appellant respectfully disagrees with this assessment. The use of Hong was not an attempt to substitute using the art of Hong. Obviously, the Hong reference does not provide for any teaching of neural networks, a key component of Appellant's invention. Instead, as indicated above, Hong was provided to the Examiner to provide evidence that particular aspects of Appellant's invention such as nanoconductors attracted to a connection gap (rather than electrodes) has been

demonstrated. Hong was <u>not</u> submitted to the Examiner as evidence of a working example of Appellant's invention, but was provided to the Examiner to provide examples of test data pictures and test documentation related to Appellant's invention, namely nanoconductors attracted to the connection gap, rather than the electrodes. Thus, the Appellant is not providing evidence of the presence or absence of working examples, but instead offers evidence of important aspects of Appellant's invention via the Hong reference, which was provided to the Examiner because the Examiner had previously cast doubt on the viability of Appellant's invention.

In this regard, Appellant cites Section (a)(4) of 35 U.S.C. §112, 37 C.F.R. § 1.105 Requirements for information, which indicates that "any reply to a requirement for information pursuant to this section that states either that the information required to be submitted is unknown to or is not readily available to the party or parties from which it was requested may be accepted as a complete reply". The information requested by the Examiner is not readily available to the Application. Information related to particular aspects of Appellant's invention, however, is available, and this is why the Hong reference was submitted to the Examiner. Appellant notes again that the Hong reference demonstrates particular aspects of Appellant's invention as indicated above, but does not demonstrate a working example of Appellant's invention. Instead, Appellant's invention was constructively reduced to practice via the filing of Appellant's specification and claims.

The filing of a patent application serves as conception and constructive reduction to practice of the subject matter described in the application. Thus, Appellant need not provide evidence of either conception or actual reduction to practice when relying on the content of the patent application. *Hyatt v. Boone*, 146 F.3d 1348, 1352, 47 USPQ2d 1128, 1130 (Fed. Cir. 1998). Additionally, constructive reduction to practice: "[O]ccurs upon the filing of a patent application on the claimed invention." *Brunswick Corp. v. U.S.*, 34 Fed. Cl. 532, 584 (1995).

Appellant's filed application IS evidence of the constructive reduction to practice and thus the request by the Examiner for "test data" is improper, particularly due to the Examiner's apparent misunderstanding of Appellant's

invention. The Appellant again submits that the Hong reference demonstrates merely particular aspects of Appellant's invention, but not Appellant's invention itself. It is unnecessary to provide "test data" because Appellant's invention provides numerous examples of aspects of Appellant's invention and also because Appellant's invention was constructively reduced to practice. Additionally, it is improper for the Examiner to request such "test data" given the Examiner's apparent lack of understanding of Appellant's invention. It would be improper to provide "test data" because Appellant does not currently have access to such test data. Appellant can merely provide examples of references that demonstrate particular aspects of Appellant's invention but can also point to examples of patented neural network / nanotechnology patents as indicated via the IX. EVIDENCE APPENDIX herein.

D. The nature of the invention: The Examiner argued that the essence of the invention requires special skills on a multidisciplinary basis exhibited by the University Group ... and thus such is skill not to be found in one of ordinary skill in the art.

The Appellant respectfully disagrees with this assessment. It seems that the Examiner is asserting that it is impossible for individuals to be multidisciplinary. For example, under this argument, a physicist could not work with a mechanical engineer and/or a chemist to develop innovative technologies and patent such multidisciplinary innovations. The Appellant does not understand how this argument applies to the Hong reference, which again is not even prior art with respect to Appellant's invention.

One of ordinary skill in the art could replicate Appellant's invention based solely on Appellant's specification. FIG. 9-10, for example, of Appellant's specification, provides a simple methodology that one skilled in the art could follow to replicate Appellant's invention. Many of the other figures in Appellant's specification also provide such guidance along with Appellant's complete written description including background, etc. Appellant's invention is not a complicated concept, but actually constitutes a simple and elegant solution to problems outlined in Appellant's background section of Appellant's specification. One skilled in the art

reviewing Appellant's specification and claims would clearly see this. Thus, the nature of Appellant's invention is quite simple and does not require more than a basic knowledge of dielectrophoresis and neural networks and the proper machinery to replicate Appellant's invention. Appellant submits that one skilled in the art could actually replicate Appellant's invention based on Appellant's claims alone, without event referring to Appellant's specification. Appellant, however, does recommend that one skilled in the art review Appellant's specification for additional assistance.

In any event, Appellant disagrees that the essence of Appellant's invention requires special skills on a multidisciplinary basis exhibited by the University Group ... and the Examiner's argument that such is skill not to be found in one of ordinary skill in the art. What does this mean? How is this argument by the Examiner a demonstration of the nature of Appellant's invention. It is almost as if the Examiner is arguing that Appellant's invention is too cross-disciplinary to be patented.

Appellant additionally notes that the Hong reference does not provide for any teaching of neural networks. Thus, in this sense, Hong and the University Group represent an improper reference upon which to base an argument that that the essence of Appellant's invention requires special skills on a multidisciplinary basis exhibited by the University Group etc. In fact, the University Group was not really crossing disciplines as they were all working in the field of dielectrophoresis, which involves the use of nanoconductors and a dielectric medium. The lack of teaching by Hong of neural networks renders Hong irrelevant in light of determining the nature and skill of those skilled in the art re: neural networks, of which Appellant's invention is based.

E. The state of the prior art: The Examiner argued that the University Group sets forth such considerations on page S665, column 2, lines 3-24 at a point in time one year after Appellant's filing of the application. The Examiner asserted that it would be significantly beyond one of ordinary skill in the art to replicate the invention by filling in using the skill of one of ordinary skill in the art at the time of the invention to create the invention represented by the specification ... the Examiner argued that the neural network could not be constructed.

The Appellant respectfully disagrees with this assessment. Examiner's argument seems to be irrelevant and that first, Hong is not a prior art reference so could not be used to point out the state of the art at the time of Appellant's invention. Second, the Examiner seems to be arguing that the University Group had such difficulty performing their specific experiment that a neural network based on Appellant's synapse could not possibly be constructed. This is an invalid argument because their difficulties are specifically related to a unique combination of nanoparticle electrode and solution material in addition to a self assembled monolayer and furthermore, the work of the University Group is in no way intended to construct a neural network. The University Group could, for example, be using latex nanoparticles and be building on the sufficient body of knowledge related to that use. However, "university groups" in general is/are almost exclusively working on new and innovative experiments and will always be encountering difficulties because of this. A "university group" in general could not even get published if they were not attempting something new and therefore difficult.

The Hong reference again, merely, demonstrates that important aspects of Appellant's invention were implemented at a point in time after Appellant's filing. The considerations mentioned by the University Group on page S665, column 2, lines 3-24 of the Hong reference merely cast doubt on only the difficult of that particular experiment, but do not provide any basis for casting doubt on Appellant's invention as indicated above.

In order to consider the state of the art, the Examiner should also consider the following patents, issued by the U.S. Patent & Trademark Office, copies of which are included herewith in Appendix IX (EVIDENCE):

U.S. Patent No.	Title
7,107,252	Pattern recognition utilizing a nanotechnology -based neural network
7,039,619	Utilized nanotechnology apparatus using a neural network, a solution and a connection gap
7,028,017	Temporal summation device utilizing nanotechnology
6,995,649	Variable resistor apparatus formed utilizing nanotechnology
6,889,216	Physical neural network design incorporating nanotechnology

These patents, also inventions of the Appellant, can assist the Examiner under this "state of the prior art" category in establishing the state of the art of nanotechnology-based neural networks and related devices and methods.

In any event, Appellant's submits that one of ordinary skill in the art could replicate Appellant's invention based solely on Appellant's specification. Appellant's invention is not a complicated concept, but actually a simple and elegant solution to problems outlined in Appellant's background section of Appellant's specification.

F. The relative skill of those in the art: The Examiner argued that such skills were not at the level of the University Group one year after filing of the application.

The Appellant respectfully disagrees with this assessment. Again, the Appellant notes that Appellant's invention is based on three primary fields, that is, micro-electronics, dielectrophoresis, and neural networks. The combination of these three basic areas is new and unobvious and it is simply incorrect to argue that there were no people skilled in these arts at the time of the filing of Appellant's invention. Regarding the University Group, the qualifications of the researches of the University Group seem more than sufficient to be able to replicate Appellant's invention given Appellant's specification and claims. For example, the University Group includes researches from the Department of Electronics and Computer Engineering (e.g., microelectronics), Institute of Quantum Information Processing and Systems (e.g., neural networks), and the Department of Chemical and Biochemical Engineering (e.g., dielectrophoresis). Thus, Appellant disagrees that such skills were not at the level of the University Group after the filing of Appellant's invention.

G. The predictability or unpredictability of the art: The Examiner argued that the issue is control of an object in a solution with physical features that is of the order of nanometers. The Examiner asserted that the University Group cites in the Hong art at page S666, column 1, lines 1-4, the systematic experimentation necessary to achieve the results noted in the Hong art. The Examiner argued that Figs. 2 and 3 of Hong cite the effects of parameter variation for Au Nano-particles.

The Examiner argued that such data represents variability in results but only addresses Au Nano-particles. The Examiner argued that the "real issue" is how does one assemble something like Fig. 3 of the Appellant specification in a repeatable manner ... or just once such that it represents a connection network that can be used in assembling a neural network.

The Appellant respectfully disagrees with this assessment. The Examiner's arguments are completely irrelevant and are based on an inaccurate assumption on how computational systems can be built or used. In essence, the Examiner is arguing that Appellant's connection network cannot be assembled in a repeatable manner. The Appellant notes that Appellant's connections described in Appellant's specification is by its very nature stochastic and therefore not necessarily repeatable at specific time scales. This is whole point of Appellant's invention. The Appellant can point, for example, to the field of molecular biology where there are mechanisms that are entirely reliable yet built on stochastic processes. For example, the movement of molecules in the inter-cellular fluid of a cell. In addition, Appellant's specification [see paragraph 00104] which indicates that the nanoconnections may or may not be arranged in an orderly array pattern between the input and output electrodes. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

Thus, repeatability of the formations of connections at a specific time scale is not needed. Rather, it is the statistics of many nanoconductors being attracted to a variety of gaps over time that results in a reliable system. The Appellant urges the Examiner to look into the reliability of a biolological synapse and points out that the

Examiner's reliability of through processes is built on a foundation of unreliable synaptic elements (as is everyone's).

Appellant further notes that the Hong reference is entitled "Controllable Capture of Au Nano-particles by using dielectrophoresis". The term "controllable" is synonymous with predictability.

H. The breadth of the claims: The Examiner argued that claims broadly cover an electromechanical neural network formed by a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes.

The Appellant actually agrees with this assessment. Appellant's claims do cover a type of electromechanical neural network. Appellant does not understand why this is a problem and also how and why this relates to Hong reference.

Back to the Hong reference. The Hong reference is but one example of a particular aspect of Appellant's invention (i.e., attracting nanoconductors to a connection gap rather than to the electrodes). The Hong reference does not teach a neural network and also post-dates Appellant's application. There are, however, other examples of neural networks based on nanotechnology that have been patented and which therefore, given such patenting, exhibit a constructive reduction to practice. The following are examples of U.S. patents issued after the filing date of Appellant's application, which utilize nanotechnology and nanoconductors in the context of neural networks:

U.S. Patent No.	Title
7,107,252	Pattern recognition utilizing a nanotechnology -based neural network
7,039,619	Utilized nanotechnology apparatus using a neural network, a solution and a connection gap
7,028,017	Temporal summation device utilizing nanotechnology
6,995,649	Variable resistor apparatus formed utilizing nanotechnology
6,889,216	Physical neural network design incorporating nanotechnology

Such references are included herewith in IX. EVIDENCE APPENDIX. would probably be more appropriate for the Examiner to consider than the Hong reference if the Examiner desires to review examples of artificial and physical neural networks and related components based on nanotechnology that have already been

patented. Appellant's present invention represents a non-obvious and novel variation to the neural network devices/techniques described in the aforementioned U.S. patents.

Based on the foregoing, the Appellant submits that the claims 24, 40 and 42 (and hence all of claims 24-44) comply with the enablement requirement of 35 U.S.C. 112, first paragraph in that Appellant's disclosure provides sufficient examples of Appellant's invention along with a sufficient description, which one skilled in the art can readily replicate. The Hong reference in this regard and as explained above is irrelevant. The Appellant respectfully requests withdrawal of the rejection to claims 24, 40 and 42 under 35 U.S.C. 112, first paragraph.

APPELLANT'S ARGUMENTS REGARDING ISSUE #2 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Claims 24-44 comply with the written description requirement of 35 U.S.C., first paragraph.

Claims 24-44 were rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The Examiner argued that the claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The Examiner argued that the above discussion regarding enablement applies. The Examiner asserted that the specification starting on page 10 – page 93 does not set forth the "how to make" electromechanical neural network formed by a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes. The Examiner argued that while control of electrical characteristics might be helpful, one has to first have the basis of the neural network and asserted that the Appellant has not disclosed and not provided this in the request made in the Office Action dated May 22, 2007.

The Appellant respectfully disagrees with this assessment. The Appellant submits that that the claim(s) do contain subject matter was described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor, at the time the application was filed, had possession of the claimed invention. A neural network is by definition a network of neural nodes. If the Appellant can demonstrate a neural node and mechanisms for the connections of nodes, then the Appellant has demonstrated that a neural network can be built.

There are numerous instances in Appellant's specification where the electromechanical neural network formed by a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes are shown and described. Appellant submits that for a full appreciation and understanding of Appellant's electromechanical neural network constituting a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes, one must review the entire specification in its entirety.

Examples of the electromechanical neural network constituting a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes are shown throughout Appellant's specification. FIG. 37, for example shows an electromechanical neural network 3700 which includes a pre-synaptic electrode 3702 and a post-synaptic electrode 3704 and nanoconnections 3708, wherein as explained in paragraph [00304] – [00305], nanoconnections 3708 are located within a connection gap 3710, which is illustrated generally by a dashed line in FIG. 37. Nanoconnections 3708 generally are formed as a plurality of interconnected nanoconnections. The, connection gap 3710 can be filled with a solvent or solution (e.g., a liquid, gel, etc), which as indicated throughout Appellant's specification is preferably dielectric, i.e. dielectric solution. For example, paragraph [00189] of Appellant's specification indicates the use of a layer of a layer of nano-conductor/dielectric solution between two electrodes. Appellant's specification also indicates that such a dielectric solution can be a liquid, gel, etc. Paragraph [00305] indicates that individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire(s), a nanotube(s), nanoparticles(s), or any other molecular structures (e.g., molecules). Nanoconnections 3708 can also be constituted as a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections 3708 can be formed from a plurality of interconnected nanoparticles (i.e. molecules), depending upon design considerations. These are but a few examples among examples in Appellant's specification of the enablement of an electromechanical neural network constituting a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes are shown throughout Appellant's specification.

Another example of an electromechanical neural network constituting a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes is shown in FIG. 38, which is similar to that shown in FIG. 38 but also shows circuitry 3718 which is used to control the neural network made up of the electrodes 3702, 3704, nanoconnections 3708 (i.e., connected nanoconductors) and the dielectric solution, etc.

FIGS. 14-17 and paragraphs [00184] – [00208] also illustrate and describe an electromechanical neural network constituting a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes.

The use of a dielectric solution in association with Appellant's nanoconductors suspended in the dielectric solution between two electrodes is clearly set forth in Appellant's specification, including the "how to make" the electromechanical neural network formed by a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes. FIGS. 9-10 of Appellant's specification including the related description clearly explain how this is accomplished and actually indicate a method for using such a dielectric solution.

The aforementioned examples of Appellant's electromechanical neural network are but a few of many that are shown and described in Appellant's specification. Thus, the Examiner is incorrect in asserting that Appellant's specification does not set forth the "how to make" electromechanical neural network formed by a plurality of nanoconductors suspended in a liquid dielectric solution between two electrodes. The Appellant is not merely describing the control of electrical characteristics that are helpful in forming Appellant's electromechanical neural network. Appellant's specification and hence claims, demonstrate the basis of Appellant's electromechanical neural network. In this regard, Appellant has satisfied the request made in the Office Action dated May 22, 2007, because Appellant's specification previously satisfied the disclosure of not only an electromechanical neural network, but also the disclosure of nanoconductors disposed in a connection gap filled with a dielectric solution between electrodes.

APPELLANT'S ARGUMENTS REGARDING ISSUE #3 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

> **Appellant's concept of a liquid dielectric solution comprising a mixture of a plurality of nanoconductors and a liquid dielectric solvent wherein a plurality of nanoconductors are free to move about in a dielectric solution and such a solution is disposed between two electrodes is NOT anticipated by Paul M. Adriani and Alice P. Gast in the article entitled "Electric-field-induced aggregation in dilute colloidal suspensions" published in 1990 by the Faraday Discussions of the Chemical Society, hereinafter referred to as the Adriani reference.**

The Examiner argued that Appellant's concept of a liquid dielectric solution comprising a mixture of a plurality of nanoconductors and a liquid dielectric solvent wherein a plurality of nanoconductors are free to move about in a dielectric solution and such solution is disposed between two electrodes is anticipated by Paul M. Adriani and Alice P. Gast in the article entitled "Electric-field-induced aggregation in dilute colloidal suspensions" (hereinafter referred to as "Adriani" or the "Adriani" reference) published in 1990 by the Faraday Discussions of the Chemical Society. In support of this argument, the Examiner cited abstract of the Adriani reference as follows:

> Electric-field induced chain formation in dilute, non-aqueous suspensions of sterically stabilized, 1 um poly (methyl methacrylate) (PMMMA) lattices are investigated. Optical microscopy and digital image analysis provide the chain-length distribution. We find that the particles carry a charge sufficient to inhibit field-induced aggregation. Equilibrium predictions of chain aggregation incorporating a screened Coulombic repulsion and field-induced dipole attraction agree with experimental observations near the onset of aggregation; chain formation becomes diffusion limited above the threshold field strength.

The Appellant respectfully disagrees with this assessment. First, the Appellant points out that the Adriani reference is dealing with a thousand nanometer scale lattice. The Examiner has noted that nanotechnology refers to dimensions that are less than 1000 nanometers. Therefore, according to the Examiner's own definition of nanotechnology, Adriani is not likely applicable as a nanotechnology reference. Second, Adriani does not teach neural networks. Adriani does not provide for any teaching of neural networks and neural network components such as adaptive synaptic components and so forth. Third, Adriani does not discuss circuits and/or any microelectronic devices and systems. Adriani is purely focused on the theoretical and experimental validation of the formation of

particle chains. For Adriani to anticipate Appellant's invention, there must at a minimum be some teaching of neural networks. This is not the case. Thus, Appellant submits that Adriani is not a proper reference with respect to Appellant's invention.

Appellant further notes that Adriani provides no teaching, suggestion or disclosure of a synaptic element, a synapse, or an adaptive synaptic element. The Examiner has cited Adriani but has not identified which specific components of Adriani constitute a synapse and an adaptive synaptic element. Again, the Examiner has not identified which aspects of Adriani constitute a neural network. Where is a neural network shown in Adriani?

Additionally, Adriani provides no disclosure, suggestion, teaching of an adaptive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution. Adriani also does not provide a teaching, suggestion or disclosure of a plurality of interconnected nanoconnections associated with said adaptive synaptic element, said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution. There is also not a teaching in Adriani of a voltage mechanism for applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors, thereby creating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field

across said connection gap. The use of colloidal suspensions and chains and electroheological fluids by Adriani is not a demonstration of the creation of a dielectrophoretic force or the strengthening or weakening of nanoconductors as taught by Appellant's claims and specification with respect to an <u>actual</u> neural network.

The Appellant therefore submits that the Adriani reference does not anticipate each and every claim limitation of Appellant's claims and as such, Adriani does not anticipate each and every claim limitation of Appellant's claims. "Anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration." *W.L. Gore & Associates v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303, 313 (Fed. Cir. 1983) (citing *Soundscriber Corp. v. United States*, 360 F.2d 954, 960, 148 USPQ 298, 301 (Ct. Cl.), *adopted*, 149 USPQ 640 (Ct. Cl. 1966)), *cert. denied*, 469 U.S. 851 (1984). Thus, to anticipate the Appellants' claims, the Adriani reference must disclose each element of the respective claims that they are being recited for. "There must be no difference between the claimed invention and the reference disclosure, as viewed by a person of ordinary skill in the field of the invention." *Scripps Clinic & Research Foundation v. Genentech, Inc.*, 927 F.2d 1565, 18 USPQ 2d 1001, 1010 (Fed. Cir. 1991).

To overcome the anticipation rejection, the Appellants need only demonstrate that not all elements of a *prima facie* case of anticipation have been met, *i. e.*, show that the Adriani reference fails disclose every element in each of the Appellants' claims associated with the relevant reference used for their rejection. "If the examination at the initial state does not produce a prima face case of unpatentability, then without more the Appellant is entitled to grant of the patent." *In re Oetiker*, 977 F.2d 1443, 24 USPQ 2d 1443, 1444 (Fed. Cir. 1992).

Thus, because the Adriani reference does not anticipate each and every claim limitation of Appellant's claims 24-44, the Examiner's anticipation rejection based on Adriani has been traversed. The Appellant respectfully requests withdrawal of the rejection to claims 24-44 based on anticipation by Adriani.

APPELLANT'S ARGUMENTS REGARDING ISSUE #4 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Appellant has defined the term "nanotechnology" in Appellant's specification.

The Examiner argued that the Appellant has not defined the term Nanotechnology. The Examiner referred to the web @ www.answers.com/nanotechnology, and indicated that the following definition was obtained:

Nanotechnology: the science and technology of building devices, such as electronic circuits, from single atoms and molecules.

The Examiner further indicated that from the Nanotechnology web site created by Dr. Ralph Merkle, the statement is made that the "word nanotechnology has become very popular and is used to describe many types of research where characteristic dimensions are less than about 1,000 nanometers" (micron range). http:www.zyvex.com/nano/

The Examiner argued that the Appellant has not defined "nanotechnology" related to a specific numeric scale. The Examiner noted that the Appellant has made the following size comparison @ specification, page 6, lines 15-18:

> Microelectrical nano-size components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, microinstruments, micro-optics, and the like.

The Examiner further asserted that this definition is entirely consistent with the above cited definitions/intent.

The Appellant respectfully disagrees with this assessment. That is, the Appellant has in fact defined nanotechnology in the specification. For example, paragraphs [0016] – [0017] of Appellant's specification teach the following:

> The term "Nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

In general, nanotechnology presents a solution to the problems faced in the rapid pace of computer chip design in recent years. According to Moore's law, the number of switches that can be produced on a computer chip has doubled every 18 months. Chips now can hold millions of transistors. It is, however, becoming increasingly difficult to increase the number of elements on a chip utilizing existing technologies. At the present rate, in the next few years the theoretical limit of silicon-based chips will have been attained. Because the number of elements and components that can be manufactured on a chip determines the data storage and processing capabilities of microchips, new technologies are required for the development of higher perform ance chips.

The Appellant further indicates the following at paragraphs [0020] – [0021]:

Integrated circuits and electrical components thereof, which can be produced at a molecular and nanometer scale, include devices such as carbon nanotubes and nanowires, which essentially are nanoscale conductors ("nanoconductors"). Nanoconductors are tiny conductive tubes (i.e., hollow) or wires (i.e., solid) with a very small size scale (e.g., 0.7 to 300 nanometers in diameter and up to 1mm in length). Their structure and fabrication have been widely reported and are well known in the art. Carbon nanotubes, for example, exhibit a unique atomic arrangement, and possess useful physical properties such as one-dimensional electrical behavior, quantum conductance, and bal listic electron transport.

Carbon nanotubes are among the smallest dimensioned nanotube materials with a generally high aspect ratio and small diameter. High-quality single-walled carbon nanotubes can be grown as randomly oriented, needle-like or spaghetti-like tangled tubules. They can be grown by a number of fabrication methods, including chemical vapor deposition (CVD), laser ablation or electric arc growth.

Thus, contrary to the Examiner's assertion that the Appellant has not defined the term Nanotechnology, there are numerous examples in Appellant's specification of what is meant by the term "nanotechnology" and also examples of size, shape, and types of components (e.g., nanoconductors) that constitute nanotechnology. In general, the term nanotechnology, as taught by Appellant's specification, refers to nanometer scale devices, components and processing techniques. Thus, Appellant's specification teaches that the use of nanometer scale components and smaller, such as nanotubes, nanowires, and so forth constitutes nanotechnology.

APPELLANT'S ARGUMENTS REGARDING ISSUE #5 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Appellant has defined the term "solvent" in Appellant's specification.

The Examiner asserted that related to terminology, Appellant refers to solvent in the generic sense in the specification, and cites page 25, ¶ 0099, that includes a condition of suspension.

The Appellant respectfully disagrees with this assessment. Appellant has not referred to the term "solvent" in a vacuum. Paragraph [0099] does not refer to a solvent in a generic sense, but defines a solvent by way of an example as follows:

> A connection network can be constructed as follows. Initially, a voltage can be applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture can be composed of a variety of materials or substances. The only general requirement in constructing such a connection network is that the conducting wires should be suspended in the solvent and/or dissolved or in a suspension, but free to move about. Additionally, the electrical conductance of the substance should generally be less than the electrical conductance of the suspended conducting nanowire(s) and/or other nanoparticle(s). The viscosity of the substance should not be too much so that the conducting nanowire(s) and/or other nanoparticles(s) cannot move when an electric field is applied.

This is consistent with what is meant by a solvent. Additionally, paragraph [00117] of Appellant's specification indicates the following:

> As illustrated at block 904, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent" and also a suspension.

Appellant is uncertain based on the Examiner's final office action and previous office actions, why the definition "solvent" is even an issue to the Examiner. The term is defined in Appellant's specification and is also known in the art. It is clear that a solvent is essentially a component, such as a liquid, that dissolves a solid, liquid, or gaseous solute, resulting in a solution.

APPELLANT'S ARGUMENTS REGARDING ISSUE #6 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Appellant has defined the term "liquid dielectric solution" in Appellant's specification.

The Examiner asserted that related to terminology, Appellant refers to a liquid dielectric solution without any explicit definition of dielectric. The Examiner asserted that the term dielectric means, to one of ordinary skill in the art, a non-conducting or insulating substance which resists passage of electric current, allowing electrostatic induction to act across it. The Examiner argued that the liquid dielectric solution will inherently have an electric conductance that is less than that when the subject solution has conducting material suspended in it such as the claimed nanoconductors.

Appellant disagrees with this assessment and notes that the term "liquid dielectric solution" is adequately disclosed and defined in Appellant's specification. For example, paragraphs [00117] – [00118] discuss a fluid/liquid in the context of a solvent comprising a volatile <u>liquid</u> that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution can be sandwiched between wafers of silicon or other materials…and If the <u>fluid</u> has a melting point that is approximately at operating temperature, then the viscosity of the <u>fluid</u> could be controlled easily. Thus, a solution is formed based on the use of a solvent that is a liquid. Appellant's specification then describes, for example, the use of a dielectric solvent, such as in paragraph [00183] of Appellant's specification, which indicates a nano-conductor/<u>dielectric</u> solvent mixture, which has been described previously herein.

Paragraph [00185] of Appellant's specification describes forming a gap 1402 between two plates P1 and P2 covered with electrodes, filled with a <u>solution</u> of nano-conductors and a <u>dielectric</u> solvent, it can be appreciated that connections can easily form between every input and every output by aligning vertically from one input electrode to a perpendicular output electrode. Paragraph [00187] describes

two basic chip structures, an input layer 1606 and an output layer 1604, each sandwiched over a gap 1602 filled with a nanoconductor/dielectric solvent mixture. Paragraph [00189] describes a total chip comprising two electrode arrays aligned perpendicular to each other, with a layer of nano-conductor/dielectric solution between the two. Paragraph [00193] describes prepared nanoconductors, such as, for example, nanotubes and/or nanowires, which can be simply mixed with a dielectric solvent. A micro-drop of the solution can thereafter be placed between the electrode arrays. Paragraph [00199] indicates that a solution can be provided which is prepared from a plurality of nanoconductors and a dielectric solvent. Paragraph [00212] of Appellant's specification indicates that when particles in a dielectric solution are exposed to an electric field (i.e., AC or DC), the particles align with the field. As the particles align, the resistance between the respective electrodes decreases. Paragraph [00216] indicates, connections between neurons 1910, 1920, 1930, 1940, 1950, and 1960 may be formed as nanoconductor(s) suspended within a dielectric solvent or solution. Paragraph [00245] describes a strong nano-connection that can result in higher van-der-wall attractions and a corresponding heightened resistance to dissolution within the dielectric medium. Paragraph [00304] describes, for example, the connection gap 3710 filled with a solvent or solution (e.g., a liquid, gel, etc). These are but a few examples of a liquid dielectric solution. Appellant notes that simply using the term "dielectric" is sufficient enough for one skilled in the art. The word "dielectric" on its face is a common term of art, and is therefore sufficient. Appellant's specification clearly indicates that the resulting solution can be a liquid dielectric solution, etc. One skilled in the art would clearly understand this from a complete reading of Appellant's specification and claims.

The term dielectric speaks for itself. The Appellant partially agrees with the Examiner's definition of a dielectric, which is consistent with Appellant's specification and claims. That is, the conductance of the dielectric solution may be a highly nonlinear function of the concentration of the materials suspended in it. That is to say the conductance of the dielectric solution and a low concentration of suspended materials could be equivalent to the conductance of the dielectric solution itself.

APPELLANT'S ARGUMENTS REGARDING ISSUE #7 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

> **Whether the prior art cited by the Examiner conclusively establishes that the invention of the Appellant will not function as a neural network.**

The Examiner cited the following on page 20, line 1 of Adriana as follows: "Particles have aligned dipoles will aggregate into chains."

The Examiner then argued that Mehrotra et al. in Elements of Artificial Neural Networks cites the nature of neural networks to include a feed forward neural network in Figure 1.15 on page 20; the adaptive linear element of Figure 2.8 with weight adjustments into a summation circuit with a training algorithm identified in Figure 2.9 on page 59. The Examiner argued that Mehrotra, among others, assert neural networks with layers of nodes feeding with a plurality of connections into a plurality of nodes at the next layer.

The Examiner then cited Therese C. Jordon et al (hereinafter referred to as the "Jordon" reference or simply "Jordon") writing in 1989 in the IEEE, Entitled "Electrorheology" to cite a graphic illustration of dipole arrangement in the presence of an electric field in figure 16 on page 867 which was copied from an article published in 1978 by H.A. Pohl, entitled: "Dielectrophoresis: The behavior of neural matter in non-uniform fields." [Appellant notes that the Examiner may have misquoted the title of the Pohl reference. Appellant asserts that the words "neural natter" in the title above cited by the Examiner, should actually be been "neutral matter". Thus Pohl has nothing to do with "neural" aspects.].

The Examiner argued that such arrangements follow dipole to dipole aligned to the field between the electrodes. The Examiner asserted that there is no evidence of dipoles forming nodes and dipoles crossing from one chain to other chains as is required in the formation of neural networks. Further, in the cited Coulombic repulsion, the Examiner argued that such repulsion will prevent extension of dipoles. Additionally, the Examiner argued there is no formation of chains to form weights to adjust values at a given node.

The Examiner argued that the evidence shown in the cited papers demonstrates that the concept disclosed by the Appellant and cited below will not function as a neural network:

> An electromechanical neural network system based on nanotechnology, comprising: an adaptive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution; a plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution; and a voltage mechanism for applying an electric field across said connection gap whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors, thereby aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

The Examiner argued that Adriani and Jordan describe liquid state anticipated equivalents of Appellant's electromechanical-based liquid state machine ... albeit without a neural network. The Examiner asserted that given that the dipoles align as described by Adriani et al. and Jordan et al. into chains (and as demonstrated by them), a neural network will not form nor will a trained neural network form ... the Examiner asserted that one simply has chains of varying lengths which by Mehrotra and others is not a neural network. The Examiner therefore argued that the prior art conclusively establishes that the invention of the Appellant will not function as a neural network.

The Appellant respectfully disagrees with this assessment and first asks why is the Examiner arguing that Adriani and Jordan et al. describe liquid state anticipated equivalents of Appellant's electromechanical-based *liquid state machine*? Appellant is <u>not</u> claiming in this invention, an electromechanical-based <u>liquid state machine</u>. Appellant is instead claiming a <u>neural network</u>. The Appellant does not understand why the Examiner asserts that Appellant teaches an electromechanical-based liquid state machine in Appellant's specification/claims.

Second, the Examiner's arguments that "there is no evidence of dipoles forming nodes and dipoles crossing from one chain to other chains" as is required in the formation of neural networks is completely irrelevant as the Appellant has not claimed this feature. Instead, the Appellant's specification and claims demonstrate how an electric field can be used to induce a dipole in a nanoconductor which can be attracted to an electrode gap and thus provide the basis for a synaptic element and these synapses can be used to connect neural nodes, which can be constructed by, for example, microelectronic circuits. It seems from the sentence above by the Examiner that the Examiner has not read and does not understand the nature of Appellant's invention. The Examiner seems to be arguing that there is no evidence that dipoles form nodes. Appellant is not claiming this, which is why the Examiner's argument is irrelevant.

Regarding the combination of Jordon and Adriani as argued by the Examiner, Appellant again notes that the Examiner has clearly misunderstood the nature of Appellant's invention. That is, the aggregation of nanoconductors whether in chains or not are used for the construction of a synapse, which is an important component in the construction of a neural network, but is not in and of itself a neural network. Jordon does not provide for any teaching of a neural network and neural network components such as adaptive synaptic elements.

Regarding the Mehrotra reference, the Appellant notes that while Mehrotra teaches principals of neural networks, the Mehrotra reference does not provide for any teaching whatsoever of nanotechnology and physical neural networks or dielectrophoresis, and so on. Mehrotra is simply a textbook that generally teaches mathematical models and principals of neural networks. In other words, Mehrotra relates to software-based neural networks, which are common in the art. Actual physical neural networks, on the other hand, such as Appellant's invention, are not taught or disclosed by Mehrotra and in fact are not common in the art, because to date, such devices as explained in Appellant's background section, have been difficult to produce. The closest Mehrotra comes to teaching physical neural networks is biological networks, but provides no teaching whatsoever of artificial physical neural networks such as that taught by Appellant's invention. That is the entire point behind Appellant's development of an actual synapse and hence

electromechanical neural network. Appellant's background section in paragraphs [009] - [0015] clearly defines the problems inherent with software based neural networks such as Mehrotra and also the problems with physical neural networks to date:

> Neural networks that have been developed to date are largely software-based. A true neural network (e.g., the human brain) is massively parallel (and therefore very fast computationally) and very adaptable. For example, half of a human brain can suffer a lesion early in its development and not seriously affect its performance. Software simulations are slow because during the learning phase a standard computer must serially calculate connection strengths. When the networks get larger (and therefore more powerful and useful), the computational time becomes enormous.
>
> For example, networks with 10,000 connections can easily overwhelm a computer. In comparison, the human brain has about 100 billion neurons, each of which can be connected to about 5,000 other neurons. On the other hand, if a network is trained to perform a specific task, perhaps taking many days or months to train, the final useful result can be built or "downloaded" onto a piece of hardware and also mass-produced. Because most problems requiring complex pattern recognition are highly specific, networks are task-specific. Thus, users usually provide their own, task -specific training data.
>
> A number of software simulations of neural networks have been developed. Because software simulations are performed on conventional sequential computers, however, they do not take advantage of the inherent parallelism of neural network architectures. Consequently, they are relatively slow. One frequently used measurement of the speed of a neural network processor is the number of interconnections it can perform per second.
>
> For example, the fastest software simulations available can perform up to approximately 18 million interconnects per second. Such speeds, however, currently require expensive super computers to achieve. Even so, approximately 18 million interconnects per second is still too slow to perform many classes of pattern classification tasks in real time. These include radar target classifications, sonar target classification, automatic speaker identification, automatic speech recognition, electro-cardiogram analysis, etc.
>
> The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions and/or billions.
>
> Due to the difficulties in constructing such highly interconnected processors, currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and configured to implement one particular neural network architecture and are not easily, if at all, reconfigurable in implementing different architectures. A true physical neural network chip, with the learning abilities and connectivity of a biological network, has not yet been designed and successfully implemented.
>
> The problem with a pure hardware implementation of a neural network utilizing existing technology is the inability to physically form a great number of connections and neurons. On-chip learning can exist, but the size of the network is limited by digital processing methods and associated electronic circuitry. One of the difficulties in creating true physical neural networks lies in the highly complex manner in which a physical neural network must be designed and constructed. The present inventor believes that solutions to creating a true physical and artificial neural network lie in the use of nanotechnology and the implementation of a novel form of variable connections.

The Mehrotra reference does not provide any hint, suggestion or teaching of artificial <u>physical</u> neural networks and particularly those that are electromechanical in nature. The Mehrotra reference merely provides a teaching of software neural network models (i.e., algorithms and hence, software). For example, Pg. 57 of Mehrotra refers to various neural network <u>algorithms</u>, but provides no teaching whatsoever of an actual physical neural network. For example, there is no teaching by Mehrotra of physical neural networks made of physical components such as resistors, capacitors, transistors and so on. There is no teaching in Mehrotra of nanoconnections composed of actual physical components such as wires, conductors. The feed forward neural network in Figure 1.15 on page 20 of Mehrotra and the adaptive linear element of Figure 2.8 of Mehrotra with weight adjustments into a summation circuit with a training algorithm identified in Figure 2.9 on page 59 are merely illustrations and description of artificial neural network models based on algorithms. There is no teaching here of an actual physical neural network nor any language that would suggest how one skilled in the art would modify Mehrotra to provide for an actual physical neural network device composed of physical neural network components.

Based on the fact that Mehrotra does not teach artificial physical neural networks nor nanotechnology, is not clear how the Mehrotra reference provides evidence (as argued by the Examiner) that conclusively establishes that the invention of the Appellant will not function as a neural network.

The Jordon reference also does not provide for either a teaching of neural networks nor fundamental principles of Appellant's invention. Jordon does not provide for any teaching whatsoever of neural networks. The dipole arrangement in the presence of an electric field in figure 16 on page 867 of Jordon, cited by the Examiner which was copied from an article published in 1978 by H.A. Pohl, entitled: "Dielectrophoresis: The behavior of <u>neutral</u> matter in non-uniform fields" is not a teaching of <u>neural</u> networks, adaptive synaptic elements, pre and post synaptic electrodes, and so on. The Examiner has merely referred to this particular article cited on page 867 of Jordon, without explicitly demonstrating which aspects of this citation constitute specific components of Appellant's invention. How does the Pohl

reference, for example, constitute a teaching of neural networks? How is this a teaching of an adaptive synaptic component? Where in such a citation is a disclosure or teaching of an adaptive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution. How does Jordon and the H.A. Pohl article entitled: "Dielectrophoresis: The behavior of neural matter in non-uniform fields" demonstrate a plurality of interconnected nanoconnections associated with said adaptive synaptic element, said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution. Additionally, how does Jordon and the H.A. Pohl article entitled: "Dielectrophoresis: The behavior of neural matter in non-uniform fields" demonstrate a voltage mechanism for applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors, thereby creating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

The Examiner has cited Jordon and the H.A. Pohl, entitled: "Dielectrophoresis: The behavior of neural matter in non-uniform fields" but has not adequately explained how this provides evidence (as argued by the Examiner) that conclusively establishes that the invention of the Appellant will not function as a neural network, when in fact, Jordon and the cited H.A. Pohl article do not teach

or disclose a neural network (i.e., again, the H.A. Pohl reference refers to “neutral matter” NOT neural matter”).

Based on the foregoing, the Appellant submits that the prior art references cited by the Examiner do not conclusively establish that the invention of the Appellant will not function as a neural network, particularly when two of the references – Adriani and Jordon – do not provide for a teaching of neural networks, and the third reference – Mehrotra – relates merely to a general teaching of software neural networks than an actual teaching of physical artificial neural networks. Therefore, the Appellant requests that the Examiner withdraw his assertions that the prior art references cited by the Examiner conclusively establish that the invention of the Appellant will not function as a neural network.

APPELLANT'S ARGUMENTS REGARDING ISSUE #8 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Whether claims 24-44 lack patentable utility under 35 U.S.C. § 101.

In the final office action dated August 30, 2007, the Examiner argued that the claimed invention of the Appellant lacks patentable utility under 35 U.S.C. 101. In support of this argument, the Examiner cited paragraph ¶ 8 of the Examiner's final office action. The Examiner argued that the neural network that is claimed cannot develop and asserted that the "whatever" network that may develop, cannot function as a neural network because it is not a neural network ... the Examiner again asserted that "chains" are not neural networks.

Appellant respectfully disagrees with this assessment, and again notes that the Examiner seems to misunderstand the nature of Appellant's invention. The Appellant is not claiming that chains are neural networks. Instead, as Appellant has repeatedly pointed out herein, the dipole induced aggregation of nanoconductors in a connection gap which may or may not form into chains form the basis of an adaptive synaptic element, which is an important component of a neural network, but which of course does not constitute a neural network in its entirety. In this regard, the Examiner has failed to demonstrate a lack of patentable utility under 35 U.S.C. 101, but has instead demonstrated that the Examiner does not seem to understand the nature of Appellant's invention.

APPELLANT'S ARGUMENTS REGARDING ISSUE #9 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Whether the Appellant has disclosed the practical application for the invention and whether the rejection to claims 24-44 under 35 U.S.C. § 112 should be withdrawn because the Appellant has NOT failed as a matter of law to satisfy 35 U.S.C. § 101.

In the final office action dated August 30, 2007, the Examiner rejected claims 24-44 under 35 USC 112, first paragraph, arguing that current case law (and accordingly, the MPEP) require such a rejection if a 101 rejection is given because when Appellant has not in fact disclosed the practical application for the invention, as a matter of law there is no way Appellant could have disclosed how to practice the undisclosed practical application. The Examiner cited the MPEP as follows:

> ("The how to use prong of section 112 incorporates as a matter of law the requirement of 35 U.S.C. 101 that the specification discloses as a matter of fact a practical utility for the invention ... <u>If the application fails as a matter of fact to satisfy 35 U.S.C. 101, then the application also fails as a matter of law to enable one of ordinary skill in the art to use the invention under 35. U.S.C. § 112.</u>"); In re Kirk, '376 F.2d 936, 942, 153 USIPQ 48, 53 (CCPA 1967) ("Necessarily, compliance with § 112 requires a description of how to use presently useful inventions, otherwise an Appellant would anomalously be required to teach how to use a useless invention."). See, MPEP 21107.01 (IV), quoting in Re Kirk (emphasis added).

The Examiner rejected claims 24-44 on this basis. Appellant's invention does provide patentable utility. The nanometer-scale physical neural network taught by Appellant's claims is enabled by Appellant's specification, which provides numerous examples of neural network and nanotechnological components and devices in the context and as a basis for forming and operating and training Appellant's neural network. The Examiner has asserted that Appellant's invention fails to satisfy 35 U.S.C. 101 but has not established that this is the case. The Examiner's assertion is just that – an assertion – and the Examiner has failed to conclusively establish that Appellant's invention fails to satisfy 35 U.S.C. 101.

Appellant's specification in fact indicates that the physical neural network of Appellant's invention would be much faster than any present software-based neural network solutions, thereby providing a solution to the problems inherent with software-based neural networks and the few artificial physical neural networks that have been implemented by others in the past. These are considerations to be taken into consideration in establishing the utility of Appellant's invention, in

addition to the actual description of the various embodiments disclosed in Appellant's specification.

The fundamental concept of Appellant's electromechanical neural network is remarkably simple and is disclosed throughout Appellant's specification. When particles in a dielectric solution are exposed to an electric field (i.e., AC or DC), the particles align with the field. As the particles align, the resistance between the respective electrodes decreases. The connection becomes stable once the electric field is removed. As the strength or frequency of the applied electric field is increased, the connections become increasingly aligned and the resistance further decreases. By applying a perpendicular electric field, for example, one can also decrease the strength of the connections. Such connections can be utilized as "synapses" in a physical neural network chip and the result is neural network chip – a fully adaptable, high-density neural network chip. Appellant's FIGS. 14-18, for example, illustrates example of such neural network chips and related components. It should be understood, of course, that such figures should be read in the context of the entire specification.

Appellant's specification provides many examples of utility. For example, FIGS. 14-18 of Appellant's specification, for example, describe a chip-implementation of Appellant's invention. This constitutes one example a practical application for practicing Appellant's invention. Other examples include the methodology of FIGS. 9-10 and the neural network structure of FIGS. 37-38.

The Examiner should note that it would take over a million modern computers to simulate the basic neural activity that is ongoing in a human brain, and that a physical neural network such as that disclosed by Appellant could perform the equivalent neural computations in vastly less space and energy, approaching that of the human brain. There is incredible utility in making something that can function at one billion times the energy requirements of current systems.

APPELLANT'S ARGUMENTS REGARDING ISSUE #10 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Whether claims 24-44 fail to identify an invention (neural network) that can be evaluated under the conditions of novelty or nonobviousness with respect to 35 U.S.C. § 102 or § 103.

In the final office action dated August 30, 2007, the Examiner argued that claims 24-44 fail to identify an invention (neural network) that can be evaluated under the conditions of novelty or nonobviousness. The Examiner asserted that the approach taken using nanoconductors fails to produce a neural network, and that the claims as written have no basis in reality and cannot be evaluated because the invention does not and cannot exist. In support of this argument, the Examiner cited ¶ 5 of the final office action dated August 30, 2007. The Examiner argued that if the Appellant is not claiming a neural network, then the chains of Adriani and the dipoles of Jordan anticipate Appellant's invention.

Appellant respectfully disagrees with this assessment. Appellant is claiming a neural network. This is clearly set forth throughout Appellant's specification. FIGS. 1-38 and paragraphs [0001] – [00319] of Appellant's specification provide numerous examples of neural networks and a thorough teaching of neural networks based on nanotechnology. The chains of Adriani and the dipoles of Jordon cited by the Examiner clearly do not anticipate Appellant's invention because neither Adriani nor Jordon provide any disclosure of neural network and specific components such as synapses, pre and post synaptic electrodes, nor do such references teach the attraction of nanoconductors to a connection gap as part of the operation of a neural network. Adriani and Jordon also do not disclose/teach nanoconductors, which as defined by Appellant's specification are components such as nanotubes, nanowires and so forth. The Examiner has not identified such nanoconductors in the Adriani and Jordon references. The Examiner has made a broad statement that "asserted that the approach taken using nanoconductors fails to produce a neural network, and that the claims as written have no basis in reality and cannot be evaluated because the invention does not and cannot exist" but the Examiner has not established this, particularly because the Examiner refers to references such as Adriani and Jordon that as indicated herein by the Appellant, do not anticipate

Appellant's claimed invention. The Examiner seems to be trying to argue two mutually exclusive positions. First, the Examiner argued that the cited prior art references anticipate Appellant's invention and then the Examiner argued that that such references establish that Appellant's invention does not exist. Appellant has demonstrated herein that the prior art does not anticipate Appellant's invention and additionally, that such references do not demonstrate that Appellant's invention does not exist.

The Appellant submits that Appellant's specification and claims can in fact be evaluated under the conditions of novelty or nonobviousness because as demonstrated herein, Appellant's specification does satisfy 35 U.S.C. 101 as having utility and patentable subject matter, but also satisfies both the enablement requirement and written description requirements of 35 U.S.C. 112. Based on the foregoing, the Appellant submits that the Examiner was incorrect in asserting that claims 24-44 fail to identify an invention (neural network) that can be evaluated under the conditions of novelty or nonobviousness with respect to 35 U.S.C. § 102 or § 103.

APPELLANT'S ARGUMENTS REGARDING ISSUE #11 - ARGUMENTS IN SUPPORT OF PATENTABILITY OF CLAIMS:

Whether the requirement for information under 35 U.S.C. §112, 37 C.F.R. § 1.105 was proper.

Appellant submits that the Examiner's requirement for information under 35 U.S.C. 112 / 37 CFR 1.105 for information was not proper, given that Appellant's filing of Appellant's application constitutes a "constructive reduction to practice". That is a constructive reduction to practice occurred when Appellant's patent application on the claimed invention is filed. The filing of Appellant's patent application served as conception and constructive reduction to practice of the subject matter described in the application. Thus, Appellant need not provide evidence of either conception or actual reduction to practice when relying on the content of the patent application. *Hyatt v. Boone*, 146 F.3d 1348, 1352, 47 USPQ2d 1128, 1130 (Fed. Cir. 1998). Additionally, constructive reduction to practice occurred upon the filing of Appellant's patent application on the claimed invention." *Brunswick Corp. v. U.S.*, 34 Fed. Cl. 532, 584 (1995).

The Examiner seems to be shifting the issue away from the fact that a constructive reduction to practice existed at the time of filing in attempt to assert rejections under 35 U.S.C. 112 and 35 U.S.C. 101. As indicated previously, the Appellant's specification and claims satisfies the requirements of 35 U.S.C. 112 and 35 U.S.C. 101. Thus, the request for information under 35 U.S.C. 112 / 37 CFR 1.105 is improper.

The Examiner argued under 37 CFR §§ 1 1.105, the Appellant is required to submit laboratory notebooks and test data including pictures of test setup confirming the reduction to practice of the disclosed invention.

In response, the Appellant relies on MPEP 2138.05 as follows: "Reduction to practice may be an actual reduction or a constructive reduction to practice which occurs when a patent application on the claimed invention is filed. The filing of a patent application serves as conception and constructive reduction to practice of the subject matter described in the application. Thus the inventor need not provide evidence of either conception or actual reduction to practice when relying on the

content of the patent application." *Hyatt v. Boone*, 146 F.3d 1348, 1352, 47 USPQ2d 1128, 1130 (Fed. Cir. 1998).

The filing of the patent application constitutes evidence of a reduction to practice, being a constructive reduction to practice. Appellant further notes that under 37 CFR 1.105 (a)(4), a reply to a requirement for information pursuant to this section that states either that the information required to be submitted is unknown to or is not readily available to the party or parties from which it was requested may be accepted as a complete reply. Thus, the Appellant does not have laboratory notebooks and test data due. It is believed that this statement is sufficient for purposes of reply to the requirement for information under 37 CFR §§ 1.56(c) and 1.105. Such information is simply not readily available to the Examiner because it is not available to the Appellant. As indicated above, however, Appellant already constructively reduced the invention to practice at the time of the filing of the application.

Appellant further notes, however, that under 37 CFR 1.105(a)(1)(viii), the information requested by the Examiner for purposes of the requirement of information under 37 CFR 1.105 as reasonably necessary to properly examine or treat the matter can include *Technical information known to Appellant*. Technical information known to Application concerning the related art, the disclosure, the claimed subject matter, other factual information pertinent to patentability, or concerning the accuracy of the examiner's stated interpretation of such items. As such, the Appellant invites the Examiner to review the following nanotechnology-based neural network patents, which constitute technical information related to the present application, and which have been issued by the U.S. Patent & Trademark Office. These issued patents are as follows (see Appendix IX herein):

U.S. Patent No.	Title
7,107,252	Pattern recognition utilizing a nanotechnology -based neural network
7,039,619	Utilized nanotechnology apparatus using a neural network, a solution and a connection gap
7,028,017	Temporal summation device utilizing nanotechnology
6,995,649	Variable resistor apparatus formed utilizing nanotechnology
6,889,216	Physical neural network design incorporating nanotechnology

Copies of these patents are included herewith. If the Examiner is seeking technical information related to a neural network based on nanotechnology and nanoconductors in a dielectric solution, then the Appellant suggests that the Examiner review these patents as a part of his requirement for information under 37 CFR §§ 1.56(c) and 1.105, because these patents constitute technical information related to the present application and also because such patents demonstrate that other types of neural networks based on nanotechnology have been issued patents by the U.S. Patent & Trademark Office, thereby also establishing the state of the art of nanotechnology-based neural networks and related devices and methods. Thus, the requirement for information under 37 CFR §§ 1.56(c) and 1.105 is satisfied in that information concerning test data is not readily available to the Appellant for the Examiner and in the alternative, the above referenced related issued patents satisfy the request for information 37 CFR 1.105(a)(1)(viii).

SUMMARY OF ARGUMENTS AND CONCLUSION

The appealed claims comply with the enablement requirement and the description requirement of 35 U.S.C. 112, first paragraph. Appellant's concept of a liquid dielectric solution comprising a mixture of a plurality of nanoconductors and a liquid dielectric solvent wherein a plurality of nanoconductors are free to move about in a dielectric solution and such a solution is disposed between two electrodes is <u>not</u> anticipated by Paul M. Adriani and Alice P. Gast in the article entitled "Electric-field-induced aggregation in dilute colloidal suspensions" published in 1990 by the Faraday Discussions of the Chemical Society, hereinafter referred to as the Adriani reference. The Appellant has defined the term "nanotechnology" in Appellant's specification. The Appellant has defined the term "solvent" in Appellant's specification. The Appellant has defined the term "liquid dielectric solution" in Appellant's specification. The prior art cited by the Examiner does not conclusively establish that the invention of the Appellant will not function as a neural network. The appealed claims do not lack patentable utility under 35 U.S.C. § 101. The Appellant has in fact disclosed the practical application for the invention and the rejection to the appealed claims under 35 U.S.C. § 112 should be withdrawn because the Appellant has NOT failed as a matter of law to satisfy 35 U.S.C. § 101. The appealed claims do not fail to identify an invention (neural network) that can be evaluated under the conditions of novelty or nonobviousness with respect to 35 U.S.C. § 102 or § 103. The requirement for information under 35 U.S.C. §112, 37 C.F.R. § 1.105 was not proper given proper constructive reduction to practice at the time of filing of Appellant's invention.

Appellant respectfully submits that his arguments as well as the specification and prosecution record support that the appealed claims are allowable.

Appellants now respectfully request that the Board to reverse the rejections of Appellant's claims and instruct the Examiner to allow such claims.

Respectfully submitted,

Kermit Lopez

Dated: January 28, 2008

Kermit Lopez
Attorney for Appellant
Registration No. 41,953
ORTIZ & LOPEZ, PLLC
P.O. Box 4484
Albuquerque, NM 87196-4484
Tel. 505-314-1312
Email: klopez@olpatentlaw.com

VIII. APPENDIX

The following Appendix (VIII) provides a listing of the appealed claims as amended in the amendment/reply dated July 23, 2007:

24. An electromechanical neural network system based on nanotechnology, comprising:

an adaptive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said liquid dielectric solution possesses an electrical conductance that is less than an electrical conductance of said plurality of nanoconductors suspended in said liquid dielectric solution;

a plurality of interconnected nanoconnections associated with said adaptive synaptic element, said plurality of interconnected nanoconnections comprising said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution; and

a voltage mechanism for applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors, thereby creating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

25. The system of claim 24 further comprising:

a plurality of synapses associated with said adaptive synaptic element, wherein said plurality of synapses comprises said plurality of interconnected

nanoconnections of said adaptive synaptic element and wherein each synapse among said plurality of synapse adapts to a voltage independent of voltage polarization.

26. The system of claim 24 wherein each nanoconductor among said plurality of nanoconductors comprises a carbon nanotube.

27. The system of claim 24 wherein each nanoconductor among said plurality of nanoconductors comprises gold.

28. The system of claim 24 wherein each nanoconductor among said plurality of nanoconductors comprises latex.

29. The system of claim 24 wherein each nanoconductor among said plurality of nanoconductors comprises DNA.

30. The system of claim 24 wherein each nanoconductor among said plurality of nanoconductors comprises silicon.

31. The system of claim 25 wherein said voltage mechanism for applying an electric field across said connection gap applies a voltage across a space occupied by said liquid dielectric solution to configure and arrange said adaptive synaptic element, said voltage comprising a DC voltage or an AC voltage that when applied across said liquid dielectric solution gradually forms said interconnected nanoconnections of said plurality of interconnected nanoconnections in said liquid dielectric solution within said connection gap between said at least one pre-synaptic electrode and said at least one post-synaptic electrode.

32. The system of claim 24 further comprising a feedback mechanism connected to and associated with said adaptive synaptic element, wherein said feedback mechanism together with said adaptive synaptic element comprise a multi-layer, feed-forward network.

33. The system of claim 24 further comprising a feedback mechanism connection to said adaptive synaptic element that provides for a Hebbian synapse modification that permits said adaptive synaptic element to function as a recurrent and highly interconnected network.

34. The system of claim 24 further comprising a learning mechanism connected to said adaptive synaptic element to train said electromechanical neural network system, wherein said learning mechanism trains said electromechanical physical neural network utilizing an STDP (Spike-Timing Dependent-Plasticity) training rule.

35. The system of claim 31 wherein said voltage mechanism generates a weak alternating electric current perpendicular to said plurality of interconnected nanoconnections, which causes at least some of said plurality of interconnected nanoconnections not contributing to a desired output to weaken and eventually dissolve back into said liquid dielectric solution, thereby allowing for an increased flexibility in a continuous training of said electromechanical neural network system utilizing said training mechanism.

36. The system of claim 34 wherein at least some of said interconnected nanoconnections among said plurality of interconnected nanoconnections are weakened by increasing a temperature of said liquid dielectric solution, which causes at least some of said plurality of interconnected nanoconnections not contributing to a desired output to weaken and eventually dissolve back into said liquid dielectric solution, thereby allowing for an increased flexibility in a continuous training of said electromechanical neural network system utilizing said training mechanism.

37. The system of claim 25 further comprising:

at least one base neuron in a perpendicular array structure composed of a plurality of neural network layers coupled with said plurality of synapses comprising said plurality of interconnected nanoconnections of said adaptive synaptic element,

wherein each synapse among said plurality of synapses comprises a plurality of connections conduits separated by a particular distance wherein each connection conduit among said plurality of connection conduits is a result of said plurality of nanoconductors aligning in the presence of said electric field, wherein said electric field is generated by a temporal and sequential firing of said at least one base neuron.

38. The system of claim 24 wherein said dielectric solvent comprises a volatile liquid and further comprising an air-tight seal for confining said liquid dielectric solution within said connection gap to prevent said volatile liquid from coming into contact with air.

39. The system of claim 25 further comprising a gate located adjacent said connection gap, and insulated from electrical contact by an insulation layer, wherein said gate is connected to logic circuitry which can activate or deactivate at least one synapse among said plurality of synapses utilizing a gate voltage provided by said gate, whereby a resistance between said at least one pre-synaptic electrode and said at least one post-synaptic electrode is modifiable by aligning said plurality of nanoconductors in said liquid dielectric solution when said electric field is applied across said connection gap and comprises an alternating electric field.

40. An electromechanical neural network system based on nanotechnology, comprising:

a resistive synaptic element comprising a plurality of nanoconductors suspended and free to move about in a liquid dielectric solution located within a connection gap formed between at least one pre-synaptic electrode and at least one post-synaptic electrode, wherein said liquid dielectric solution comprises a mixture of said plurality of nanoconductors and a dielectric solvent, wherein said resistive synaptic element functions as an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode;

a plurality of interconnected nanoconnections associated with said resistive synaptic element, said plurality of interconnected nanoconnections comprising said

plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution;

a plurality of synapses associated with said resistive synaptic element, wherein said plurality of synapses comprises said plurality of interconnected nanoconnections of said resistive synaptic element and wherein each synapse among said plurality of synapses is independent of voltage polarization; and

a voltage mechanism for applying an AC electric field across said connection gap, whereby said AC electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening each nanoconnection among said plurality of interconnected nanoconnections according to an application of said AC electric field across said connection gap so that said electromechanical neural network system adapts itself to the requirements of a given situation regardless of the initial state of said plurality of interconnected nanoconnections, wherein the longer the amount of time said AC electric field is applied across said connection gap and/or the greater the frequency or amplitude of said AC electric field applied across said connection gap, the more nanoconductors among said plurality of nanoconductors align and the stronger said interconnected nanoconnections among said plurality of nanoconnections become..

41. The system of claim 40 further comprising a learning mechanism connected to said resistive synaptic element to train said electromechanical neural network system, wherein said learning mechanism trains said electromechanical physical neural network utilizing an STDP (Spike-Timing Dependent-Plasticity) training rule.

42. A method of forming an electromechanical neural network system based on nanotechnology, comprising:

providing a liquid dielectric solution comprising a mixture of a dielectric solvent and a plurality of nanoconductors, wherein each nanoconductor among said plurality of nanoconductors is suspended and free to move about in said liquid dielectric solution;

forming a connection gap between at least one pre-synaptic electrode and said at lease post-synaptic electrode;

configuring a connection network to comprise said plurality of nanoconductors suspended and free to move about in said liquid dielectric solution located within said connection gap formed between said at least one pre-synaptic electrode and said at least one post-synaptic electrode, wherein said connection network comprises an impermanent interconnect between said at least one pre-synaptic electrode and said at least one post-synaptic electrode;

configuring said connection network to comprises a plurality of interconnected nanoconnections with said connection network, wherein said plurality of interconnected nanoconnections comprise said plurality of nanoconductors in said liquid dielectric solution, said plurality of interconnected nanoconnections electrically connecting said at least one pre-synaptic electrode to said at least one post-synaptic electrode through said liquid dielectric solution and said plurality of nanoconductors disposed within said liquid dielectric solution;

providing a plurality of synapses from said connection network, wherein said plurality of synapses comprises said plurality of interconnected nanoconnections of said connection network and wherein each synapse among said plurality of synapses is independent of voltage polarization; and

applying an electric field across said connection gap, whereby said electric field induces a dipole in each nanoconductor among said plurality of nanoconductors only when said plurality of nanoconductors is located within said liquid dielectric solution, thereby generating a dielectrophoretic force attracting said plurality of nanoconductors to said connection gap and aligning said plurality of nanoconductors within said liquid dielectric solution and strengthening or weakening

each nanoconnection among said plurality of interconnected nanoconnections according to an application of said electric field across said connection gap.

43. The method of claim 42 wherein the longer the amount of time said electric field is applied across said connection gap and/or the greater the frequency or amplitude of said electric field applied across said connection gap, the more nanoconductors among said plurality of nanoconductors align and the stronger said interconnected nanoconnections among said plurality of nanoconnections become.

44. The method of claim 42 further comprising training said electromechanical neural network system utilizing an STDP (Spike-Timing Dependent-Plasticity) training rule.

IX. EVIDENCE APPENDIX

U.S. Patent No.	Title
7,107,252	Pattern recognition utilizing a nanotechnology -based neural network
7,039,619	Utilized nanotechnology apparatus using a neural network, a solution and a connection gap
7,028,017	Temporal summation device utilizing nanotechnology
6,995,649	Variable resistor apparatus formed utilizing nanotechnology
6,889,216	Physical neural network design incorporating nanotechnology

Non-Patent Reference:

N. Mureau, E. Mendoza, K. Hoettges, S.R.P. Silva and M.P. Hughes, "In-situ and real time determination of metallic and semiconducting SWNT in suspension via dielectrophoresis", Appl. Phys. Lett., 88, (2006) 243109.

US007107252B2

(12) **United States Patent**
Nugent

(10) **Patent No.: US 7,107,252 B2**
(45) **Date of Patent: Sep. 12, 2006**

(54) **PATTERN RECOGNITION UTILIZING A NANOTECHNOLOGY-BASED NEURAL NETWORK**

(75) Inventor: **Alex Nugent**, Santa Fe, NM (US)

(73) Assignee: **Knowm Tech, LLC**, Albuquerque, NM (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 3 days.

(21) Appl. No.: **11/045,564**

(22) Filed: **Jan. 31, 2005**

(65) **Prior Publication Data**

US 2005/0149464 A1 Jul. 7, 2005

Related U.S. Application Data

(63) Continuation of application No. 10/095,273, filed on Mar. 12, 2002, now Pat. No. 6,889,216.

(51) **Int. Cl.**
G06F 15/18 (2006.01)
G06E 1/00 (2006.01)

(52) **U.S. Cl.** **706/15**

(58) **Field of Classification Search** 706/15; 705/15

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

2,707,223 A		4/1955	Hollman	338/32 R
3,833,894 A		9/1974	Aviram et al.	365/151
4,802,951 A		2/1989	Clark et al.	156/630
4,926,064 A		5/1990	Tapang	706/26
4,974,146 A		11/1990	Works et al.	364/200
4,988,891 A		1/1991	Mashiko	307/201
5,315,162 A		5/1994	McHardy et al.	307/201
5,422,983 A		6/1995	Castelaz et al.	395/24
5,475,794 A		12/1995	Mashiko	395/24
5,589,692 A		12/1996	Reed	257/23
5,649,063 A		7/1997	Bose	395/22
5,670,818 A		9/1997	Forouhi et al.	257/530
5,706,404 A		1/1998	Colak	395/24
5,717,832 A		2/1998	Steimle et al.	395/24
5,761,115 A		6/1998	Kozicki et al.	365/182
5,783,840 A		7/1998	Randall et al.	257/24
5,812,993 A		9/1998	Ginosar et al.	706/26
5,896,312 A		4/1999	Kozicki et al.	365/153
5,904,545 A		5/1999	Smith et al.	438/455
5,914,893 A		6/1999	Kozicki et al.	365/107
5,951,881 A		9/1999	Rogers et al.	216/41
5,960,391 A	*	9/1999	Tateishi et al.	704/232
5,978,782 A		11/1999	Neely	706/16
6,026,358 A		2/2000	Tomabechi	704/232

(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 022 764 A1 1/2000

(Continued)

OTHER PUBLICATIONS

Quantum-Dot Array for Computation, ORNL Review Vo. 34, No. 2, 2001.*

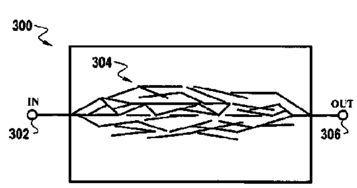
(Continued)

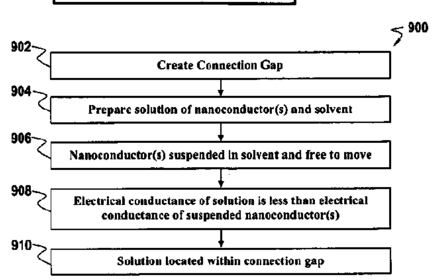
Primary Examiner—George Davis
(74) *Attorney, Agent, or Firm*—Kermit D. Lopez; Luis M. Ortiz; Ortiz & Lopez, PLLC

(57) **ABSTRACT**

A pattern recognition system, comprising a neural network formed utilizing nanotechnology and a pattern input unit, which communicates with the neural network, wherein the neural network processes data input via the pattern input unit in order to recognize data patterns thereof. Such a pattern recognition system can be implemented in the context of a speech recognition system and/or other pattern recognition systems, such as visual and/or imaging recognition systems.

20 Claims, 6 Drawing Sheets





U.S. PATENT DOCUMENTS

6,070,140 A *	5/2000	Tran	704/275
6,084,796 A	7/2000	Kozicki et al.	365/153
6,128,214 A	10/2000	Kuekes et al.	365/151
6,245,630 B1	6/2001	Ishikawa	438/393
6,248,529 B1	6/2001	Connolly	435/6
6,256,767 B1	7/2001	Kuekes et al.	716/9
6,282,530 B1	8/2001	Huang	706/41
6,294,450 B1	9/2001	Chen et al.	438/597
6,314,019 B1	11/2001	Kuekes et al.	365/151
6,330,553 B1	12/2001	Uchikawa et al.	706/2
6,335,291 B1	1/2002	Freeman	438/706
6,339,227 B1	1/2002	Ellenbogen	257/40
6,359,288 B1	3/2002	Ying et al.	257/14
6,363,369 B1	3/2002	Liaw et al.	706/15
6,383,923 B1	5/2002	Brown et al.	438/666
6,389,404 B1	5/2002	Carson et al.	706/18
6,407,443 B1	6/2002	Chen et al.	257/616
6,418,423 B1	7/2002	Kambhatla et al.	706/15
6,420,092 B1	7/2002	Yang et al.	430/311
6,422,450 B1	7/2002	Zhou et al.	228/121.85
6,423,583 B1	7/2002	Avouris et al.	438/132
6,424,961 B1	7/2002	Ayala	706/25
6,426,134 B1	7/2002	Lavin et al.	428/300.1
6,620,346 B1	9/2003	Schulz et al.	252/519.51
6,798,692 B1	9/2004	Kozicki et al.	365/174
6,855,329 B1	2/2005	Shakesheff et al.	424/409
2001/0004471 A1	6/2001	Zhang	427/372.2
2001/0023986 A1	9/2001	Mancevski	257/741
2001/0024633 A1	9/2001	Lee et al.	423/447.3
2001/0031900 A1	10/2001	Margrave et al.	570/126
2001/0041160 A1	11/2001	Margrave et al.	423/460
2001/0044114 A1	11/2001	Connolly	435/6
2002/0001905 A1	1/2002	Choi et al.	438/268
2002/0004028 A1	1/2002	Margrave et al.	423/447.3
2002/0004136 A1	1/2002	Gao et al.	428/367
2002/0030205 A1	3/2002	Varshavsky	257/208
2002/0075126 A1	6/2002	Reitz et al.	338/21
2002/0086124 A1	7/2002	Margrave et al.	428/36.9
2002/0090468 A1	7/2002	Goto et al.	427/580
2002/0102353 A1	8/2002	Mauthner et al.	427/255.28
2003/0031438 A1	2/2003	Kambe et al.	385/122
2003/0177450 A1	9/2003	Nugent	716/1
2003/0236760 A1	12/2003	Nugent	706/26
2004/0039717 A1	2/2004	Nugent	706/27
2004/0150010 A1	8/2004	Snider	257/209
2004/0153426 A1	8/2004	Nugent	706/25
2004/0162796 A1	8/2004	Nugent	706/27
2004/0193558 A1	9/2004	Nugent	706/25

FOREIGN PATENT DOCUMENTS

EP	1 046 613 A2	4/2000
EP	1 100 106 A2	5/2001
EP	1 069 206 A2	7/2001
EP	1 115 135 A1	7/2001
EP	1 134 304 A2	9/2001
RU	2071126 C1	6/1996
WO	WO 00/44094	7/2000
WO	WO 03/017282 A1	8/2001

OTHER PUBLICATIONS

Meyer et al, "Computational Neural Networks: A General Purpose Tool for Nanotechnology", 5th Forsight Conference on Molecular Nanotechnology.*

Peter Weiss, "Circuitry in a Nanowire: Novel Growth Method May Transform Chips," Science News Online, vol. 161, No. 6; Feb. 9, 2002.

Press Release, "Nanowire-based electronics and optics comes one step closer," Eureka Alert, American Chemical Society; Feb. 1, 2002.

Weeks et al., "High-pressure nanolithography using low-energy electrons from a scanning tunneling microscope," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 38-42; Dec. 12, 2001.

CMP Cientifica, "Nanotech: the tiny revolution", CMP Cientifica, Nov. 2001.

Diehl, et al., "Self-Assembled, Deterministic Carbon Nanotube Wiring Networks,"Angew. Chem. Int. Ed. 2002, 41, No. 2; Received Oct. 22, 2001.

G. Pirio, et al., "Fabrication and electrical characteristics of carbon nanotube field emission microcathodes with an integrated gate electrode," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 1-4, Oct. 2, 2001.

Leslie Smith, "An Introduction to Neural Networks," Center for Cognitive and Computational Neuroscience, Dept. of Computing & Mathematics, University of Stirling, Oct. 25, 1996; http//www.cs.stir.ac.uk/~Iss/NNIntro/InvSlides.html.

V. Derycke et al., "Carbon Nanotube Inter- and Intramolecular Logic Gates," American Chemical Society, Nano Letters, XXXX, vol. 0, No. 0, A-D.

Mark K. Anderson, "Mega Steps Toward the Nanochip," Wired News, Apr. 27, 2001.

Collins et al., "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," Science, vol. 292, pp. 706-709, Apr. 27, 2001.

Landman et al., "Metal-Semiconductor Nanocontacts: Silicon Nanowires," Physical Review Letters, vol. 85, No. 9, Aug. 28, 2000.

John G. Spooner, "Tiny tubes mean big chip advances," Cnet News.com, Tech News First, Apr. 26, 2001.

Jeong-Mi Moon et al., "High-Yield Purification Process of Singlewalled Carbon Nanotubes," J. Phys. Chem. B 2001, 105, pp. 5677-5681.

"A New Class of Nanostructure: Semiconducting Nanobelts Offer Potential for Nanosensors and Nanoelectronics," Mar. 12, 2001, http://www.sciencedaily.com/releases/2001/03/010309080953.htm.

Hermanson et al., "Dielectrophoretic Assembly of Electrically Functional Microwires from Nanoparticle Suspensions," Materials Science, vol. 294, No. 5544, Issue of Nov. 2, 2001, pp. 1082-1086.

Press Release, "Toshiba Demonstrates Operation of Single-Electron Transistor Circuit at Room Temperature," Toshiba, Jan. 10, 2001.

J. Appenzeller et al., "Optimized contact configuration for the study of transport phenomena in ropes in single-wall carbon nanotubes," Applied Physics Letters, vol. 78, No. 21, pp. 3313-3315, May 21, 2001.

David Rotman, "Molecular Memory, Replacing silicon with organic molecules could mean tiny supercomputers," Technology Review, May 2001, p. 46.

Westervelt et al., "Molecular Electronics," NSF Functional Nanostructures Grant 9871810, NSF Partnership in Nanotechnology Conference, Jan. 29-30, 2001; http://www.unix.oit.umass.edu/~nano/NewFiles/FN19_Harvard.pdf.

Niyogi et al., "Chromatographic Purification of Soluble Single-Walled Carbon Nanotubes (s-SWNTs)," J. Am. Chem. Soc 2001, 123, pp. 733-734, Received Jul. 10, 2000.

Duan et al., "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices," Nature, vol. 409, Jan. 4, 2001, pp. 67-69.

Paulson, et al., "Tunable Resistance of a Carbon Nanotube-Graphite Interface," Science, vol. 290, Dec. 1, 2000, pp. 1742-1744.

Wei et al., "Reliability and current carrying capacity of carbon nanotubes," Applied Physics Letters, vol. 79, No. 8, Aug. 20, 2001, pp. 1172-1174.

Collins et al., "Nanotubes for Electronics," Scientific American, Dec. 2000, pp. 62-69.

Avouris et al., "Carbon nanotubes: nanomechanics, manipulation, and electronic devices." Applied Surface Science 141 (1999), pp. 201-209.

Smith et al., "Electric-field assisted assembly and alignment of metallic nanowires," Applied Physics Letters, vol. 77, No. 9, Aug. 28, 2000, pp. 1399-1401.

Hone et al., "Electrical and thermal transport properties of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 666-668.

Smith et al., "Structural anisotropy of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 663-665.
Andriotis et al., "Various bonding configurations of transition-metal atoms on carbon nanotubes: Their effect on contract resistance," Applied Physics Letters, vol. 76, No. 26, Jun. 26, 2000, pp. 3890-3892.
Chen et al., "Aligning single-wall carbon nanotubes with an alternating-current electric field," Applied Physics Letters, vol. 78, No. 23, Jun. 4, 2001, pp. 3714-3716.
Bezryadin et al., "Self-assembled chains of graphitized carbon nanoparticles," Applied Physics Letters, vol. 74, No. 18, May 3, 1999, pp. 2699-2701.
Bezryadin et al., "Evolution of avalanche conducting states in electrorheological liquids," Physical Review E, vol. 59, No. 6, Jun. 1999, pp. 6896-6901.
Liu et al., "Fullerene Pipes," Science, vol. 280, May 22, 1998, pp. 1253-1255.
Yamamoto et al., "Orientation and purification of carbon nanotubes using ac electrophoresis," J. Phys. D: Appl. Phys 31 (1998) L34-L36.
Bandow et al., "Purification of Single-Wall Carbon Nanotubes by Microfiltration," J. Phys. Chem. B 1997, 101, pp. 8839-8842.
Tohji et al., "Purifying single walled nanotubes," Nature, vol. 383, Oct. 24, 1996, p. 679.
Dejan Rakovic, "Hierarchical Neural Networks and Brainwaves: Towards a Theory of Consciousness," Brain & Consciousness: Proc. ECPD Workshop (ECPD, Belgrade, 1997), pp. 189-204.
Dave Anderson & George McNeill, "Artificial Neural Networks Technology," A DACS (Data & Analysis Center for Software) State-of-the-Art Report, Contract No. F30602-89-C-0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, Aug. 20, 1992.
Greg Mitchell, "Sub-50 nm Device Fabrication Strategies," Project No. 890-00, Cornell Nanofabrication Facility, Electronics—p. 90-91, National Nanofabrication Users Network.
John-William DeClaris, "An Introduction to Neural Networks," http://www.ee.umd.edu/medlab/neural/nn1.html.
"Neural Networks," StatSoft, Inc., http://www.statsoftinc.com/textbook/stevnet.html.
Stephen Jones, "Neural Networks and the Computation Brain or Maters relating to Artificial Intelligence," The Brain Project, http://www.culture.com.au/brain_proj/neur_net.htm.
David W. Clark, "An Introduction to Neural Networks"; http://members.home.net/neuralnet/introtonn/index.htm.
"A Basic Introduction to Neural Networks"; http://blizzard.gis.uiuc.edu/htmldocs/Neural/neural.html.
Meyer et al., "Computational neural networks: a general purpose tool for nanotechnology," Abstract, 5th Foresight Conference on Molecular Nanotechnology; http://www.foresight.org/Conferences/MNT05/Abstracts/Meyeabst.html.
Saito et al., "A 1M Synapse Self-Learning Digital Neural Network Chip," ISSCC, pp. 6.5-1 to 6.5-10, IEEE 1998.
Espejo, et al., "A 16×16 Cellular Neural Network Chip for Connected Component Detection," Jun. 30, 1999; http://www.imse.cnm.csic.es/Chipcat/espejo/chip-2.pdf.
Pati et al., "Neural Networks for Tactile Perception," Systems Research Center and Dept. of Electrical Engineering, University of Maryland and U.S Naval Research Laboratory. 1987; http://www.isr.umd.edu/TechReports/ISR/1987/TR_87-123/TR_87-123.phtml.
Osamu Fujita, "Statistical estimation of the number of hidden units for feedforward neural networks," Neural Networks 11 (1998), pp. 851-859.
Abraham Harte, "Liquid Crystals Allow Large-Scale Alignment of Carbon Nanotubes," CURJ (Caltech Undergraduate Research Journal), Nov. 2001, vol. 1, No. 2, pp. 44-49.
"Quantum-Dot Arrays for Computation," ORNL Review vol. 34, No. 2, 2001, pp. 1-5 http://www.ornlgov/ORNLReview/v34_2_01/arrays.htm.
Jabri, M.A. et al., "Adaptive Analog VLSI Neural Systems," Chapman & Hall, London SE1 8HN, UK, 1996, pp. 92-95.
Lipson et al., "Automatic Design and Manufacture of Robotic Lifeforms," NATURE, vol. 406, Aug. 31, 2000, pp. 974-978.
Kunitoshi Yamamoto, et al., "Rapid Communication Orientation and Purification of Carbon Nanotubes Using AC Electrophoresis", J. Phys. D. Appl. Phys 31 (1998) L34-L36.
E.S. Snow, et al., "Random networks of carbon nanotubes as electronic material", Applied Physics Letters, vol. 82, No. 12, Mar. 31, 2003, pp. 2145-2147.
R. Martel, et al., "Ambipolar Electrical Transport in Semiconducting Single-Wall Carbon Nanotubes," Physical Review Letters, vol. 87, No. 25, Dec. 17, 2001, pp. 256805-1 to 256805-4.
S. Heinze, et al., "Carbon Nanotubes as Schottky Barrier Transistors", vol. 89, No. 10, Sep. 2, 2002, pp. 106801-1 to 106801-4.
M. Dubson, et al., "Measurement of the conductivity exponent in two-dimensional percolating networks: square lattice versus random-void continuum", Physical Review B, vol. 32, No. 11, Dec. 1, 1985, pp. 7621-7623.
D.J. Frank, et al., "Highly efficient algorithm for percolative transport studies in two dimensions", Physical Review B, vol. 37, No. 1, Jan. 1, 1988, pp. 302-307.
Uma R. Karmarkar, et al., "Mechanisms and significance of spike-timing dependent plasticity," Biol. Cybern. 87, 373-382 (2002), Jan. 28, 2002.
Uma R. Karmarkar, et al., "A Model of Spike-Timing Dependent Plasticity: One or Two Coincidence Detectors?", J. Neurophysiol, vol. 88, pp. 507-513, Jul. 2002.
M.C.W. van Rossum, et al., "Stable Hebbian Learning from Spkke-Timing-Dependent Plasticity", The Journal of Neuroscience, Dec. 1, 2003, 20(23), pp. 8812-8821.
Xiaohui Xie, et al., "Spike-based learning rules and stabilization of persistent neural activity,".
Nace L. Golding, et al., "Dendritic spikes as a mechanism for cooperative long-term potentiation", NATURE, vol. 418, Jul. 18, 2002, pp. 326-330.
Ozgur Turel, et al., "Possible nanoelectronic implementation of neuromorphic networks", Dept. o f Physics and Astronomy, Stony Brook University.
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3; Sep. 9, 2003; American Chemical Society, pp. 1379-1382.
J.M. Tour, et al., "NanoCell Electronic Memories," J.Am.Chem. Soc. 2003, 125, pp. 13279-13283.
J. Zaumseil, et al., "Three-Dimensional and Multilayer Nanostructures Formed by Nanotransfer Printing," Nano Letters, 2003, vol. 3, No. 9; Jul. 31, 2003, American Chemical Society, pp. 1223-1227.
Charles D. Schaper, "Patterned Transfer of Metallic Thin Film Nanostructures by Water-Soluble Polymer Templates," Nano Letters, 2003, vol. 3, No. 9; Jul. 26, 2003, American Chemical Society, pp. 1305-1309.
C.A. Dyke, et al., "Unbundled and Highly Functionalized Carbon Nanotubes from Aqueous Reactions," Nano Letters, 2003, vol. 3, No. 9; Aug. 19, 2003, American Chemical Society, pp. 1215-1218.
J. Chung, et al., "Nanoscale Gap Fabrication by Carbon Nanotube-Extracted Lithography (CEL)", Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003, American Chemical Society, pp. 1029-1031.
O. Harnack, et al., "Rectifying Behavior of Electrically Aligned ZnO Nanorods," Nano Letters, 2003, vol. 3, No. 8; Jun. 24, 2003, American Chemical Society, pp. 1097-1101.
M. S. Kumar, et al., "Influence of electric field type on the assembly of single walled carbon nanotubes," Chemical Physics Letters 383 (2004), Dec. 2, 2003; pp. 235-239.
S.W. Lee, et al., "Dielectrophoresis and electrohydrodynamics-mediated fluidic assembly of silicon resistors," Applied Physics Letters, vol. 83, No. 18, Nov. 3, 2003, pp. 3833-3835.
R. Krupke, et al., "Simultaneous Deposition of Metallic Bundles of Single-walled Carbon Nanotubes Using Acdielectrophoresis," Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003; American Chemical Society, pp. 1019-1023.
K. Bradley, et al., "Flexible Nanotube Electronics," Nano Letters, 2003, vol. 3, No. 10; Aug. 9, 2003, America Chemical Society, pp. 1353-1355.

T.B. Jones, "Frequency-dependent orientation of isolated particle chains," Journal of Electrostatics, 25 (1990), Elsevier Science Publishers, pp. 231-244.
L.A. Nagahara, "Directed placement of suspended carbon nanotubes for nanometers-scale assembly," Applied Physics Letters, vol. 80, No. 20; May 20, 2003; pp. 3826-3828.
A. Bezryadin, et al., "Electrostatic trapping of single conducting nanoparticles between electrodes," Applied Physics Letters, 71 (9), Sep. 1, 1997, pp. 1273-1275.
S. Suzuki, et al., "Quantitative Analysis of DNA Orientation in Stationary AC Electric Fields Using Fluorescence Anisotropy," IEEE Transactions of Industry of Industry Applications, vol. 34, No. 1; Jan./Feb. 1998, pp. 75-83.
Phaedon Avouris, "Molecular Electronics with Carbon Nanotubes," Accounts of Chemical Research, vol. 35, No. 12; Jul. 31, 2002, pp. 1025-1034.
X. Liu, et al., "Electric-Field Induced Accumulation and Alignment of Carbon Nanotubes," 2002 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp. 31-34.
R. Krupke, et al., "Contacting single bundles of carbon nanotubes with alternating electric fields," Appl. Phys. A. 76, Oct. 28, 2002, pp. 397-400.
M. Law, et al., "Photochemical Sensing of NO_2 with SnO^2 Nanoribbon Nanosensors at Room Temperature," Angew. Chem. 2002, 114, Nr. 13, pp. 2511-2514.
J. Tour, et al., "Nanocell Logic Gates for Molecular Computing," IEEE Transactions on Nanotechnology, vol. 1, No. 2, Jun. 2002, pp. 100-109.
A. Leonardi, et al., "Simulation methodology for dielectrophoresis in microelectronic Lab-on-a-chip," Modeling and Simulation of Microsystems 2002, pp. 96-99.
J. Chung, et al., "Nanoscale Gap Fabrication and Integration of Carbon Nanotubes by Micromachining," Solid-State Sensor, Actuator and Microsystems Workshop, Jun. 2-6, 2003; Hilton Head Island, South Carolina, pp. 161-164.
L. Zheng, et al., "Towards Single Molecule Manipulation with Dielectrophoresis Using Nanoelectrodes," IEEE-NANO 2003, Aug. 12-14, 2003, Moscone Convention Center, San Francisco, CA; pp. 437-440, http://ieeenano2003.arc.nasa.gov/program_contents.pdf.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3, No. 10; American Chemical Society; Sep. 8, 2003; pp. 1379-1382.
R. Krupke, "Separation of Metallic from Semiconducting Single-Walled Carbon Nanotubes," Science, vol. 301; Jul. 18, 2003; pp. 344-347.
Wolfgang Maass, "On the Relevance of Time in Neural Computation and Learning," In M. Li and A. Maruoka, editors, *Proc. of the 8th International Conference on Algorithmic Learning Theory in Sendai (Japan)*, vol. 1316 of *Lecture Notes in Computer Science*, pp. 364-388. Springer (Berlin), 1997.
Wolfgang Maass, "Noisy Spiking Neurons with Temporal Coding have more Computational Power than Sigmoidal Neurons," In M. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems*, vol. 9, pp. 211-217. MIT Press (Cambridge), 1997. (pp. 1-13, including Appendix).
L. Perrinet, et al., "Emergence of filters from natural scences in a sparse spike coding scheme," Neurocomputing, 2003, pp. 1-14, http://www.laurent.perrinet.free.fr/publi/perrinet03neurocomputing.pdf.
L. Perrinet, et al., "Coherence detection in a spiking neuron via Hebbian learning," Neurocomputing, 2002, vol. 44-46, No. C., pp. 817-822, http://www.laurent.perrinet.free.fr/publi/perrinet02.pdf.
A. Jarosz, et al., "An Introductory Note on Gaussian Correlated Random Matrix," Feb. 21, 2003, pp. 1-20 http://www.if.uj.edu.pl/pl/koloSMP/prace/rndmatrix.pdf.
K. Bradley, et al., "Influence of Mobile Ions on Nanotube Based FET Devices," Nano Letters, 2003, vol. 3, No. 5; American Chemical Society, Apr. 4, 2003; pp. 639-641.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.
Nanoparticles Get Wired, Dimes Institute, Delft University of Technology, 1997.
A. Bezryadin, *Trapping Single Particle with Nanoelectrodes*, Physics News Graphics, Sep. 1997.
Snow, et al., *Nanofabrication with Proximal Probes*, Proceedings of the IEEE, Apr. 1997.
P. O'Connor, G. Gramegna, P. Rehak, F. Corsi, C. Marzocca, *CMOS Preamplifier with High Linearity and Ultra Low Noise for X-Ray Spectroscopy*, IEEE Transactions on Nuclear Science, vol. 44, No. 3, Jun. 1997, pp. 318-325.
"Elements of Artificial Neural Networks" K. Mehrotra, C. K. Mohan, S. Ranka, 1997, MIT Press, pp. 116-135.
"Self-assembled chains of graphitized carbon nanoparticles" A. Bezryadin, R. M. Westervelt, M. Tinkham, Dec. 21, 1998.
"Collective Transport in Arrays of Small Metallic Dots" A. Alan Middleton, N.S. Wingreen, 1993, The American Physical Society, 0031-9007/93/71(19)/3198(4), pp. 3198 through 3201.

* cited by examiner

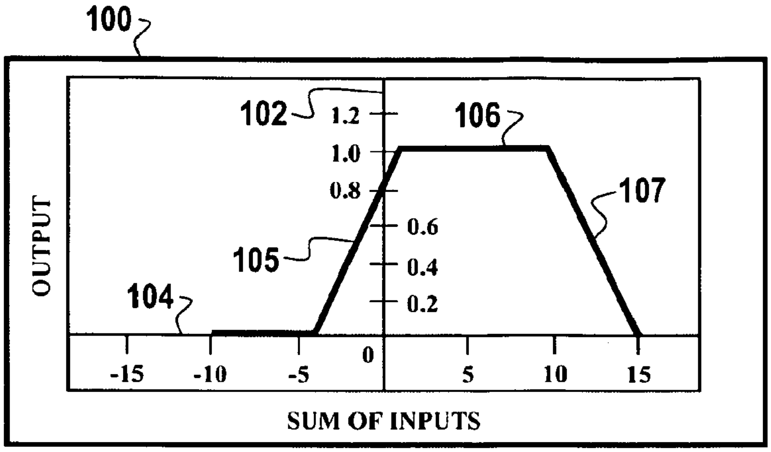


Figure 1

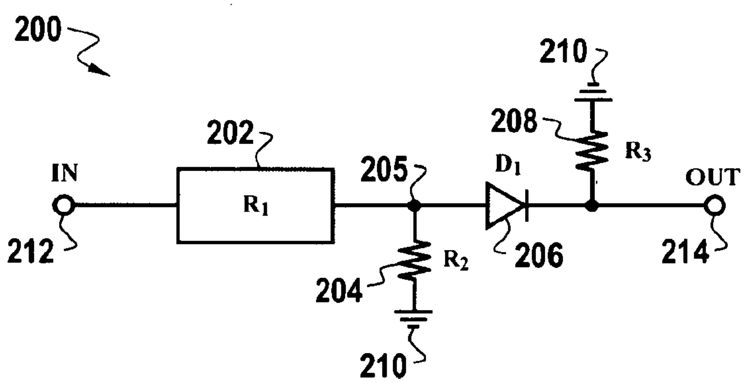


Figure 2

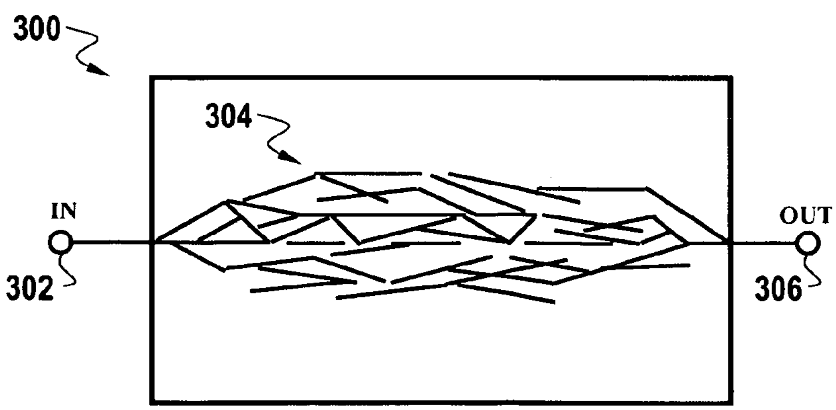


Figure 3

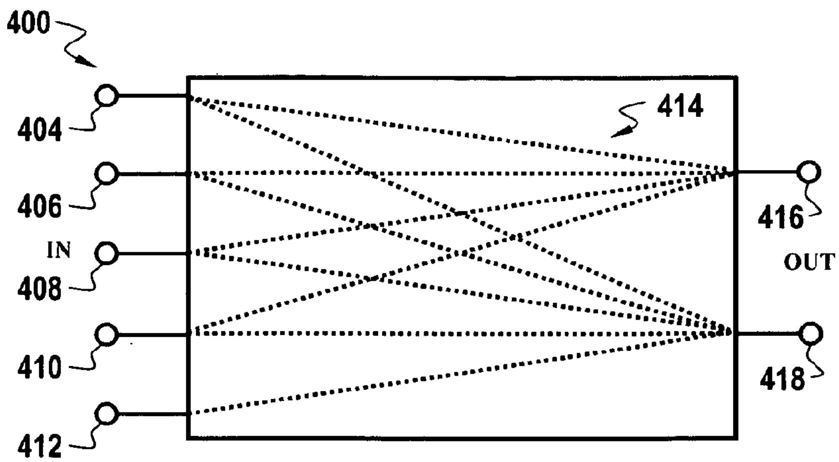


Figure 4

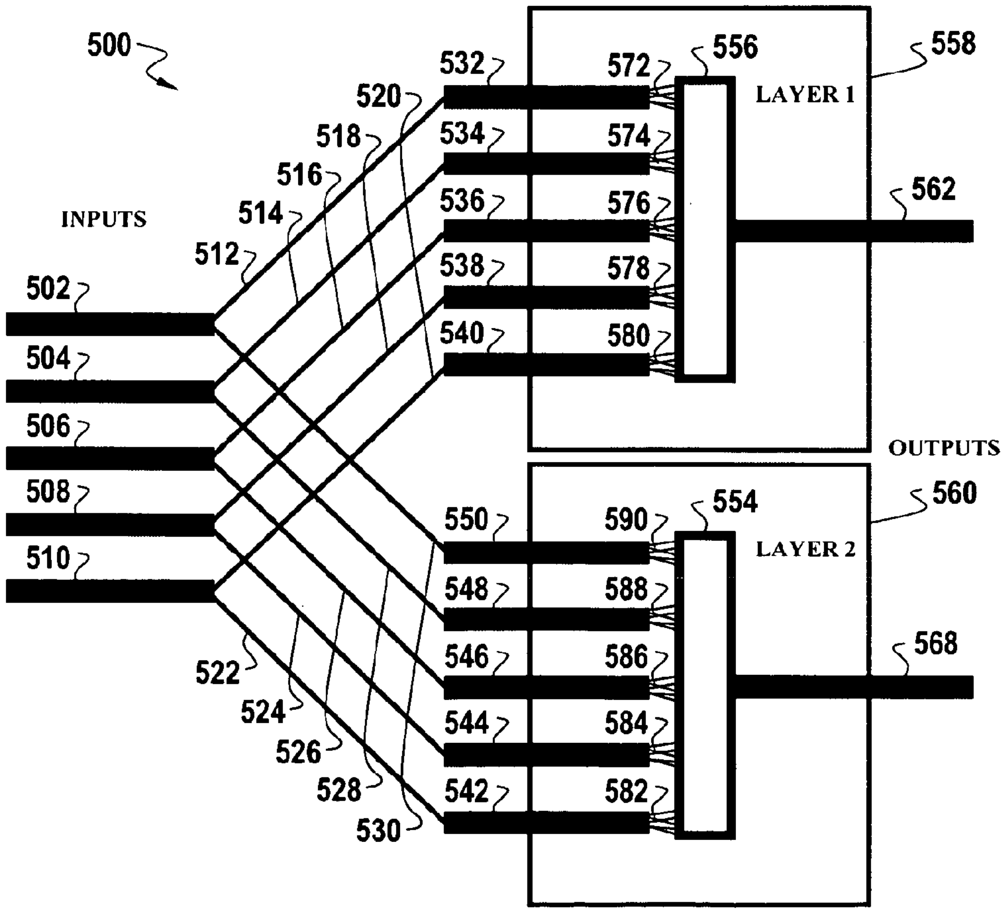


Figure 5

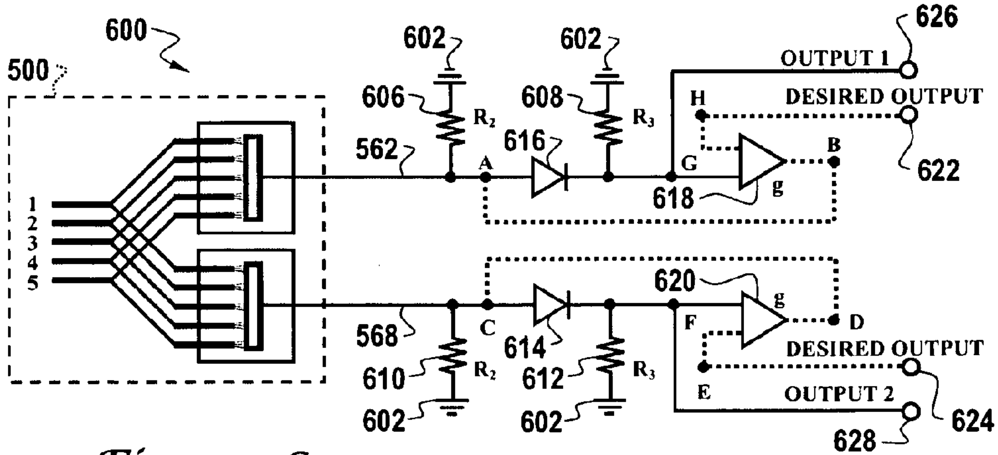
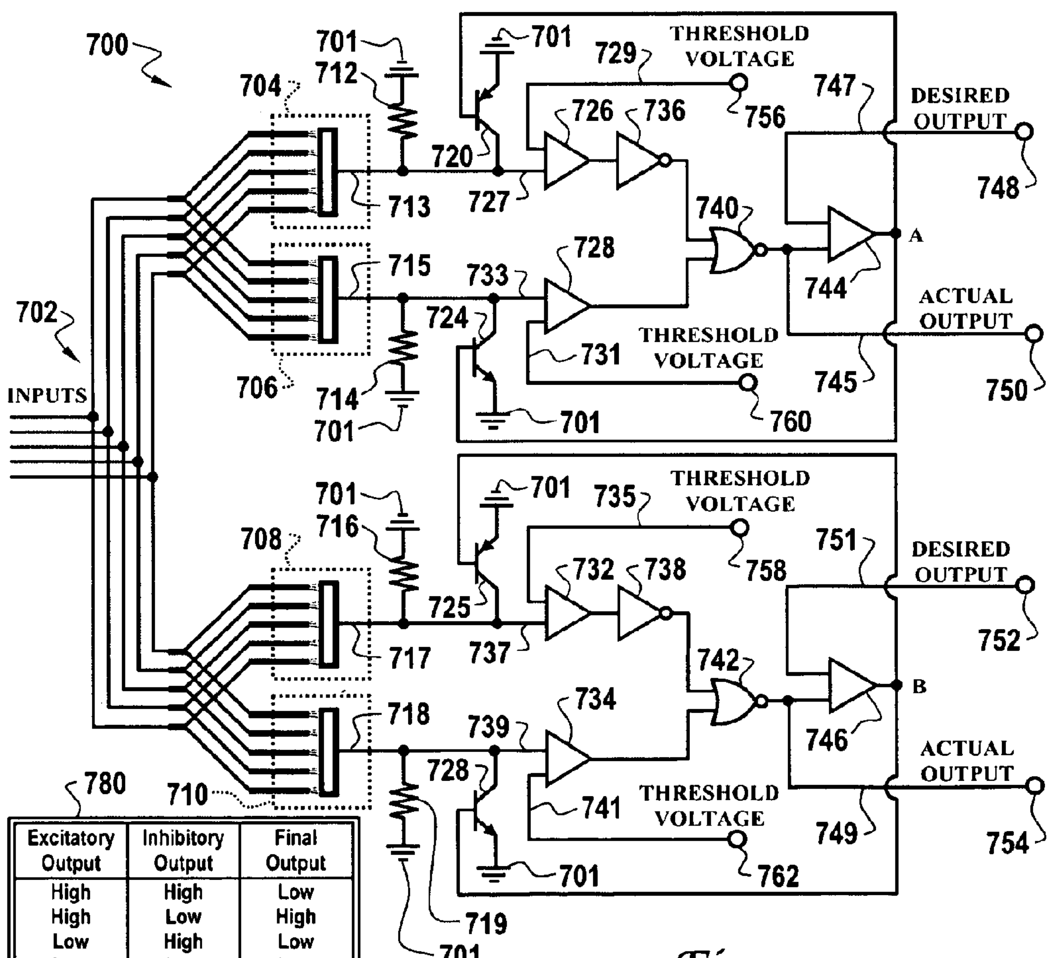


Figure 6



Excitatory Output	Inhibitory Output	Final Output
High	High	Low
High	Low	High
Low	High	Low
Low	Low	Low

Figure 7

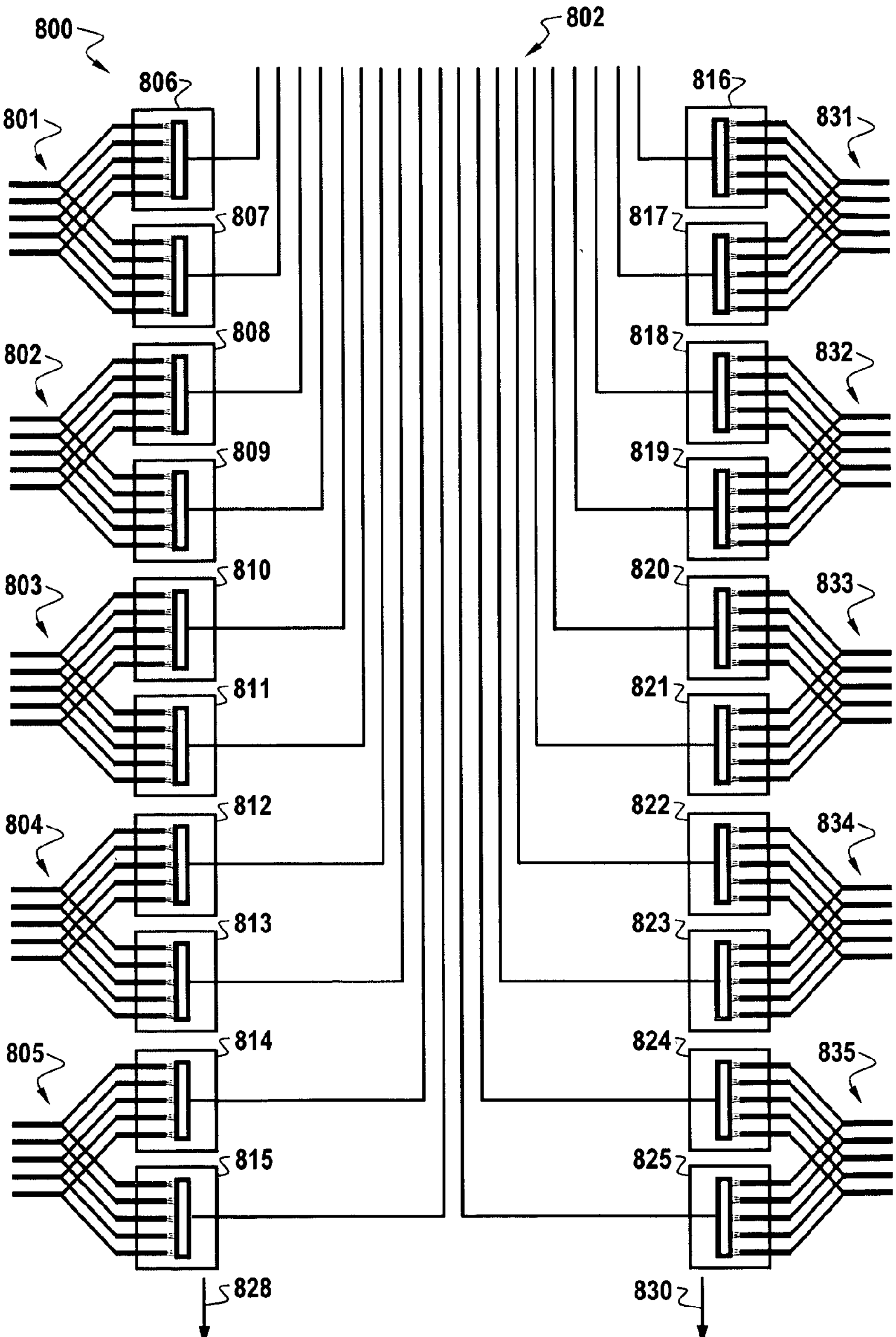

800
802
806
801
807
808
802
809
810
803
811
812
804
813
814
805
815
828
816
831
817
818
832
819
820
833
821
822
834
823
824
835
825
830

Figure 8

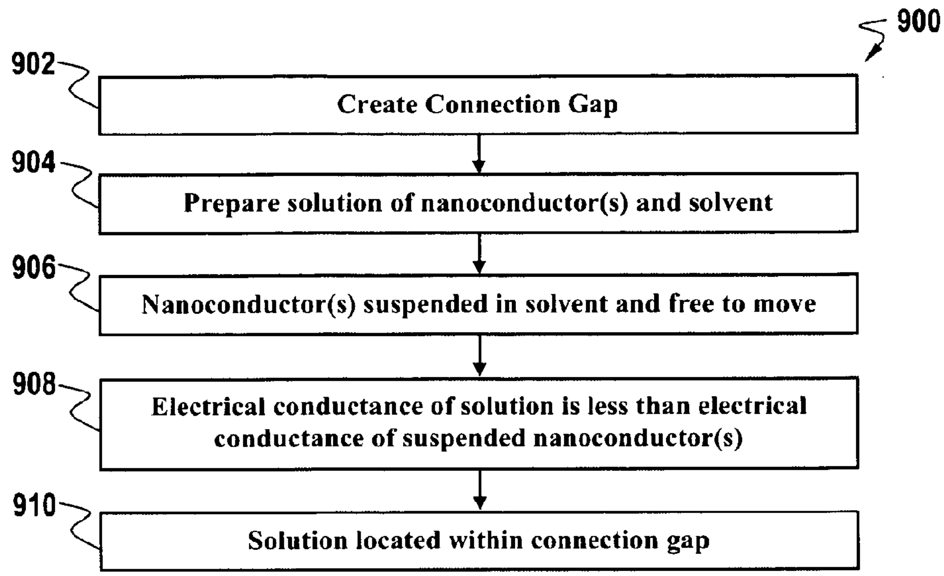


Figure 9

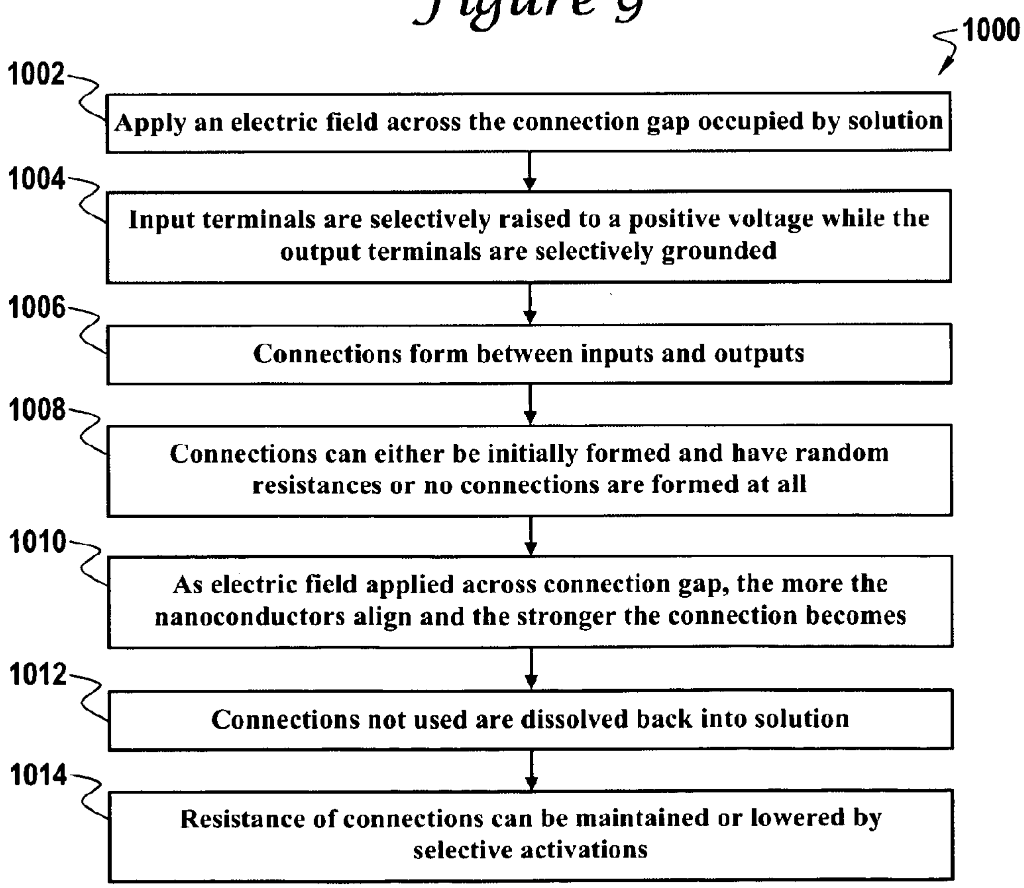


Figure 10

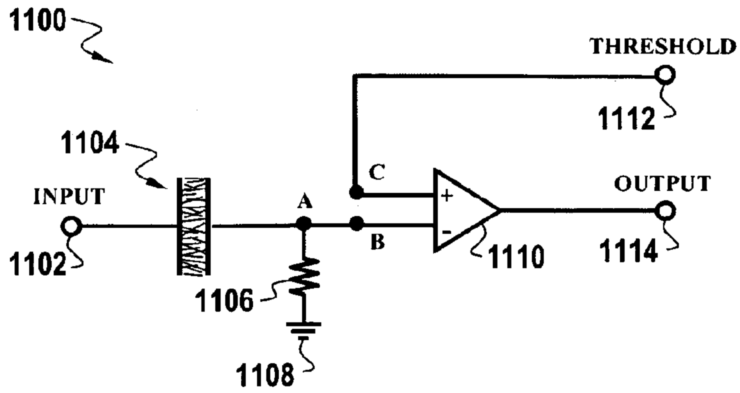


Figure 11

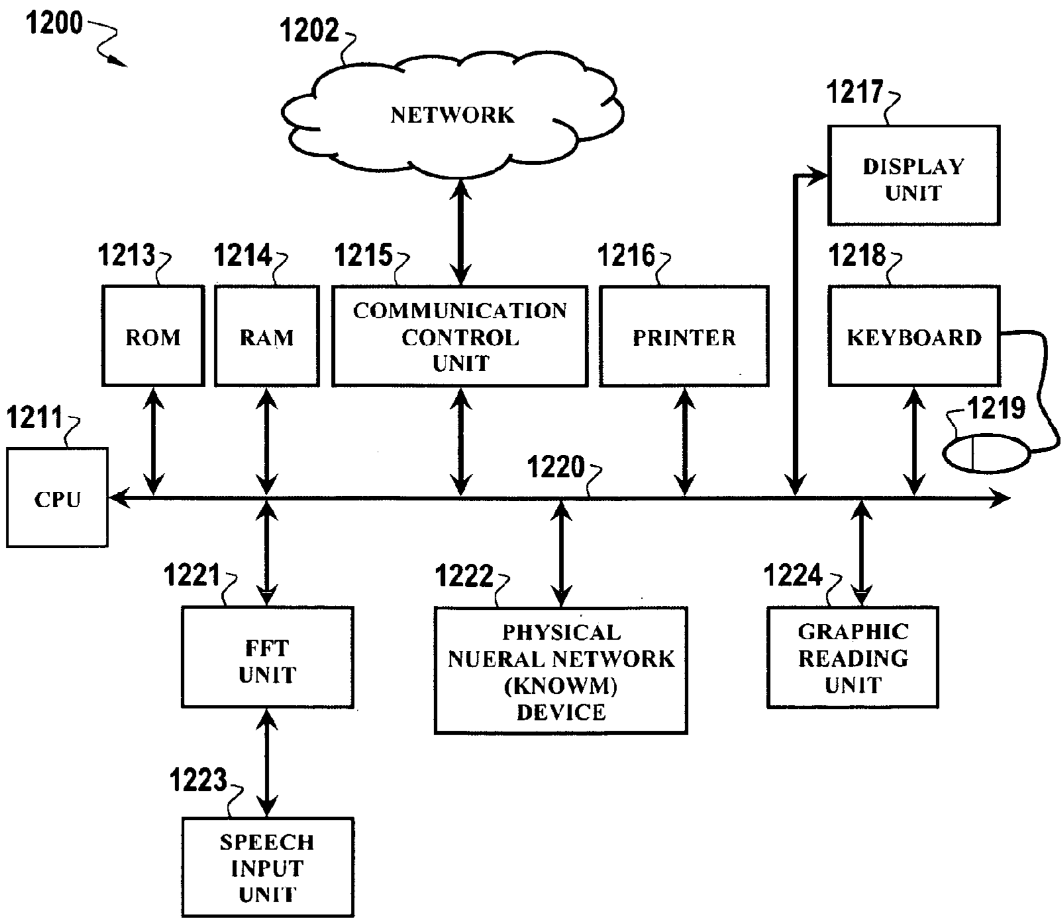


Figure 12

PATTERN RECOGNITION UTILIZING A NANOTECHNOLOGY-BASED NEURAL NETWORK

CROSS REFERENCE TO RELATED PATENT APPLICATION

This patent application is a continuation of U.S. patent application Ser. No. 10/095,273 entitled "Physical Neural Network Design Incorporating Nanotechnology," which was filed on Mar. 12, 2002, now U.S. Pat. No. 6,889,216, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

Embodiments generally relate to nanotechnology. Embodiments also relate to neural networks and neural computing systems and methods thereof. Embodiments also relate to pattern recognition devices, methods and systems, including devices that recognize speech, visual and/or imaging data.

BACKGROUND

Neural networks are computational systems that permit computers to essentially function in a manner analogous to that of the human brain. Neural networks do not utilize the traditional digital model of manipulating 0's and 1's. Instead, neural networks create connections between processing elements, which are equivalent to neurons of a human brain. Neural networks are thus based on various electronic circuits that are modeled on human nerve cells (i.e., neurons). A neural network is an information-processing network, which is inspired by the manner in which a human brain performs a particular task or function of interest.

In general, artificial neural networks are systems composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. The computational elements, or nodes, are connected via variable weights that are typically adapted during use to improve performance. Thus, in solving a problem, neural net models can explore many competing hypothesis simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights.

In a neural network, "neuron-like" nodes can output a signal based on the sum of their inputs, the output being the result of an activation function. In a neural network, there exists a plurality of connections, which are electrically coupled among a plurality of neurons. The connections serve as communication bridges among of a plurality of neurons coupled thereto. A network of such neuron-like nodes has the ability to process information in a variety of useful ways. By adjusting the connection values between neurons in a network, one can match certain inputs with desired outputs.

One does not "program" a neural network. Instead, one "teaches" a neural network by examples. Of course, there are many variations. For instance, some networks do not require examples and extract information directly from the input data. The two variations are thus called supervised and unsupervised learning. Neural networks are currently used in applications such as noise filtering, face and voice recognition and pattern recognition. Neural networks can thus be utilized as an advanced mathematical technique for processing information.

Neural networks that have been developed to date are largely software-based. The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions. Due to the difficulties in building such highly interconnected processors, the currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and built to implement one particular neural network architecture and are not easily, if at all, reconfigurable to implement different architectures. A true physical neural network (i.e., artificial neural network) chip, for example, has not yet been designed and successfully implemented.

The problem with a pure hardware implementation of a neural network with technology as it exists today, is the inability to physically form a great number of connections and neurons. On-chip learning can exist, but the size of the network would be limited by digital processing methods and associated electronic circuitry. One of the difficulties in creating true physical neural networks lies in the highly complex manner in which a physical neural network must be designed and built. It is believed that solutions to creating a true physical and artificial neural network lie in the use of nanotechnology and the implementation of analog variable connections.

The term "nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage and include devices such as nanotubes, nanowires, nanoparticles and so forth. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

Based on the foregoing, it is believed that a physical neural network which incorporates nanotechnology is a solution to the problems encountered by prior art neural network solutions. It is believed that a true physical neural network can be designed and constructed without relying on computer simulations for training, or relying on standard digital (binary) memory to store connections strengths. Additionally, such a physical neural network, if implemented properly, can be utilized for pattern recognition purposes, including speech, visual and/or imaging data.

BRIEF SUMMARY

The following summary is provided to facilitate an understanding of some of the innovative features unique to the embodiments, and is not intended to be a full description. A full appreciation of the various aspects of the embodiments can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

It is therefore another aspect of the present to provide a physical neural network, which can be formed and implemented utilizing nanotechnology.

It is another aspect of the present invention to provide for a pattern recognition system that recognizes speech, visual

data and/or imaging data and incorporates a nanotechnology-based physical neural network.

It is still another aspect of the present invention to provide a physical neural network, which can be formed from a plurality of interconnected nanoconnections or nanoconnectors.

It is yet a further aspect of the present invention to provide a physical neural network, which can be formed from a plurality of nanoconductors, such as, for example, nanowires and/or nanotubes.

It is still an additional aspect of the present invention to provide a physical neural network, which can be implemented physically in the form of a chip structure.

The above and other aspects can be achieved as is now described. A physical neural network based on nanotechnology is disclosed herein, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, connected to a plurality of interconnected nanoconnections. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes can be strengthened or weakened according to an application of an electric field. Alignment has also been observed with a magnetic field, but electric fields are generally more practical. Note that the connection network is generally associated with one or more of the neuron-like nodes.

The output signal is generally based on a threshold below which the output signal is not generated and above which the output signal is generated. The transition from zero output to high output need not necessarily be abrupt or non linear. The connection network comprises a number of layers of nanoconnections, wherein the number of layers is generally equal to a number of desired outputs from the connection network. The nanoconnections are formed without influence from disturbances resulting from other nanoconnections thereof. Such nanoconnections may be formed from an electrically conducting material. The electrically conducting material can be selected such that a dipole is induced in the electrically conducting material in the presence of an electric field. Such a nanoconnection may comprise a nanoconductor.

The connection network itself may comprise a connection network structure having a connection gap formed therein, and a solution located within the connection gap, such that the solution comprises a solvent or suspension and one or more nanoconductors. Preferably, a plurality of nanoconductors is present in the solution (i.e., mixture). Note that such a solution may comprise a liquid and/or gas. An electric field can then be applied across the connection gap to permit the alignment of one or more of the nanoconductors within the connection gap. The nanoconductors can be suspended in the solvent, or can lie at the bottom of the connection gap on the surface of the chip. Studies have shown that nanotubes can align both in the suspension and/or on the surface of the gap. The electrical conductance of the mixture is less than the electrical conductance of the nanoconductors within the solution.

The nanoconductors within the connection gap thus experience an increased alignment in accordance with an increase in the electric field applied across the connection gap. Thus, nanoconnections of the neuron-like node that are utilized most frequently by the neuron-like node become stronger with each use thereof. The nanoconnections that are utilized least frequently become increasingly weak and eventually dissolve back into the solution. The nanoconnections may or may not comprise a resistance, which can be raised or lowered by a selective activation of a nanoconnection. They can be configured as nanoconductors such as, for example, a nanotube or nanowire. An example of a nanotube, which may be implemented in accordance with the invention described herein, is a carbon nanotube, nanowire and/or other nanoparticle. Additionally, such nanoconnections may be configured as a negative connection associated with the neuron-like node.

In general, a pattern recognition system is disclosed herein, comprising a physical neural network formed utilizing nanotechnology and a pattern input unit, which communicates with the physical neural network, wherein the physical neural network processes data input via the pattern input unit in order to recognize data patterns thereof. Such a pattern recognition system can be implemented in the context of a speech recognition system and/or other pattern recognition systems, such as visual and/or imaging recognition systems.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a graph illustrating a typical activation function that can be implemented in accordance with one embodiment;

FIG. **2** illustrates a schematic diagram illustrating a diode configuration as a neuron, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanowires between two electrodes, in accordance with a preferred embodiment;

FIG. **4** illustrates a block diagram illustrating a plurality of connections between inputs and outputs of a physical neural network, in accordance with a preferred embodiment;

FIG. **5** illustrates a schematic diagram of a physical neural network that can be created without disturbances, in accordance with a preferred embodiment;

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **8** illustrates a schematic diagram of a chip layout for a connection network that may be implemented in accordance with an alternative embodiment;

FIG. **9** illustrates a flow chart of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment;

FIG. **10** illustrates a flow chart of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment;

FIG. **11** illustrates a schematic diagram of a circuit illustrating temporal summation within a neuron, in accordance with a preferred embodiment; and

FIG. **12** illustrates a block diagram illustrating a pattern recognition system, which may be implemented with a physical neural network device, in accordance with a preferred embodiment.

DETAILED DESCRIPTION

The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate one or more embodiments.

The physical neural network described and disclosed herein is different from prior art forms of neural networks in that the disclosed physical neural network does not require a computer simulation for training, nor is its architecture based on any current neural network hardware device. The design of the physical neural network described herein with respect to particular embodiments is actually quite "organic". Such a physical neural network is generally fast and adaptable, no matter how large such a physical neural network becomes. The physical neural network described herein can be referred to generically as a Knowm. The terms "physical neural network" and "Knowm" can be utilized interchangeably to refer to the same device, network, or structure.

Network orders of magnitude larger than current VSLI neural networks can be built and trained with a standard computer. One consideration for a Knowm is that it must be large enough for its inherent parallelism to shine through. Because the connection strengths of such a physical neural network are dependant on the physical movement of nano-connections thereof, the rate at which a small network can learn is generally very small and a comparable network simulation on a standard computer can be very fast. On the other hand, as the size of the network increases, the time to train the device does not change. Thus, even if the network takes a full second to change a connection value a small amount, if it does the same to a billion connections simultaneously, then its parallel nature begins to express itself.

A physical neural network (i.e., a Knowm) must have two components to function properly. First, the physical neural network must have one or more neuron-like nodes that sum a signal and output a signal based on the amount of input signal received. Such a neuron-like node is generally non-linear in its output. In other words, there should be a certain threshold for input signals, below which nothing is output and above which a constant or nearly constant output is generated or allowed to pass. This is a very basic requirement of standard software-based neural networks, and can be accomplished by an activation function. The second requirement of a physical neural network is the inclusion of a connection network composed of a plurality of interconnected connections (i.e., nanoconnections). Such a connection network is described in greater detail herein.

FIG. **1** illustrates a graph **100** illustrating a typical activation function that can be implemented in accordance with one embodiment. Note that the activation function need not be non-linear, although non-linearity is generally desired for learning complicated input-output relationships. The activation function depicted in FIG. **1** comprises a linear function, and is shown as such for general edification and illustrative purposes only. As explained previously, an activation function may also be non-linear.

As illustrated in FIG. **1**, graph **100** includes a horizontal axis **104** representing a sum of inputs, and a vertical axis **102** representing output values. A graphical line **106** indicates threshold values along a range of inputs from approximately −10 to +10 and a range of output values from approximately 0 to 1. As more neural networks (i.e., active inputs) are established, the overall output as indicated at line **105** climbs until the saturation level indicated by line **106** is attained. If a connection is not utilized, then the level of output (i.e., connection strength) begins to fade until it is revived. This phenomenon is analogous to short term memory loss of a human brain. Note that graph **100** is presented for generally illustrative and edification purposes only and is not considered a limiting feature of the embodiments.

In a Knowm network, the neuron-like node can be configured as a standard diode-based circuit, the diode being the most basic semiconductor electrical component, and the signal it sums may be a voltage. An example of such an arrangement of circuitry is illustrated in FIG. **2**, which generally illustrates a schematic diagram illustrating a diode-based configuration as a neuron **200**, in accordance with a preferred embodiment. Those skilled in the art can appreciate that the use of such a diode-based configuration is not considered a limitation of the embodiments, but merely represents one potential arrangement in which the embodiments may be implemented.

Although a diode may not necessarily be utilized, its current versus voltage characteristics are non-linear when used with associated resistors and similar to the relationship depicted in FIG. **1**. The use of a diode as a neuron is thus not a limiting feature, but is only referenced herein with respect to a preferred embodiment. The use of a diode and associated resistors with respect to a preferred embodiment simply represents one potential "neuron" implementation. Such a configuration can be said to comprise an artificial neuron. It is anticipated that other devices and components may be utilized instead of a diode to construct a physical neural network and a neuron-like node (i.e., artificial neuron), as indicated here.

Thus, neuron **200** comprises a neuron-like node that may include a diode **206**, which is labeled D_1, and a resistor **204**, which is labeled R_2. Resistor **204** is connected to a ground **210** and an input **205** of diode **206**. Additionally, a resistor **202**, which is represented as a block and labeled R_1 can be connected to input **205** of diode **206**. Block **202** includes an input **212**, which comprises an input to neuron **200**. A resistor **208**, which is labeled R_3, is also connected to an output **214** of diode **206**. Additionally, resistor **208** is coupled to ground **210**. Diode **206** in a physical neural network is analogous to a neuron of a human brain, while an associated connection formed thereof, as explained in greater detail herein, is analogous to a synapse of a human brain.

As depicted in FIG. **2**, the output **214** is determined by the connection strength of R_1 (i.e., resistor **202**). If the strength of R_1's connection increases (i.e., the resistance decreases), then the output voltage at output **214** also increases. Because diode **206** conducts essentially no current until its threshold voltage (e.g., approximately 0.6V for silicon) is attained, the output voltage will remain at zero until R_1 conducts enough current to raise the pre-diode voltage to approximately 0.6V. After 0.6V has been achieved, the output voltage at output **214** will increase linearly. Simply adding extra diodes in series or utilizing different diode types may increase the threshold voltage.

An amplifier may also be added to the output **214** of diode **206** so that the output voltage immediately saturates at the diode threshold voltage, thus resembling a step function, until a threshold value and a constant value above the threshold is attained. R_3 (i.e., resistor **208**) functions generally as a bias for diode **206** (i.e., D_1) and should generally be about 10 times larger than resistor **204** (i.e., R_2). In the circuit configuration illustrated in FIG. **2**, R_1 can actually be configured as a network of connections composed of many inter-connected conducting nanowires (i.e., see FIG. **3**). As explained previously, such connections are analogous to the synapses of a human brain.

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections **304** formed between two electrodes, in accordance with a preferred embodiment. Nanoconnections **304** (e.g., nanoconductors) depicted in FIG. **3** are generally located between input **302** and output **306**. The network of nanoconnections depicted in FIG. **3** can be implemented as a network of nanoconductors. Examples of nanoconductors include devices such as, for example, nanowires, nanotubes, and nanoparticles.

Nanoconnections **304**, which are analogous to the synapses of a human brain, are preferably composed of electrical conducting material (i.e., nanoconductors). It should be appreciated by those skilled in the art that such nanoconductors can be provided in a variety of shapes and sizes without departing from the teachings herein. For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles.

For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles. For example, each nanoconnection within a connection network may be formed from as a chain of carbon nanoparticles. In "Self-assembled chains of graphitized carbon nanoparticles" by Bezryadin et al., Applied Physics Letters, Vol. 74, No. 18, pp. 2699–2701, May 3, 1999, for example, a technique is reported, which permits the self-assembly of conducting nanoparticles into long continuous chains. Thus, nanoconductors which are utilized to form a physical neural network (i.e., Knowm) could be formed from such nanoparticles. It can be appreciated that the Bezryadin et al is referred to herein for general edification and illustrative purposes only and is not considered to limit the embodiments.

It can be appreciated that a connection network as disclosed herein may be composed from a variety of different types of nanoconductors. For example, a connection network may be formed from a plurality of nanoconductors, including nanowires, nanotubes and/or nanoparticles. Note that such nanowires, nanotubes and/or nanoparticles, along with other types of nanoconductors can be formed from materials such as carbon or silicon. For example, carbon nanotubes may comprise a type of nanotube that can be utilized in accordance with one or more embodiments.

As illustrated in FIG. **3**, nanoconnections **304** comprise a plurality of interconnected nanoconnections, which from this point forward, can be referred to generally as a "connection network." An individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire, a nanotube, nanoparticles(s), or any other nanoconducting structures. Nanoconnections **304** may comprise a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections **304** may be formed from a plurality of interconnected nanoparticles. A connection network is thus not one connection between two electrodes, but a plurality of connections between inputs and outputs. Nanotubes, nanowires, nanoparticles and/or other nanoconducting structures may be utilized, of course, to construct nanoconnections **304** between input **302** and input **306**. Although a single input **302** and a single input **306** is depicted in FIG. **3**, it can be appreciated that a plurality of inputs and a plurality of outputs may be implemented in accordance with the embodiments, rather than simply a single input **302** or a single output **306**.

FIG. **4** illustrates a block diagram illustrating a plurality of nanoconnections **414** between inputs **404**, **406**, **408**, **410**, **412** and outputs **416** and **418** of a physical neural network, in accordance with a preferred embodiment. Inputs **404**, **406**, **408**, **410**, and **412** can provide input signals to connections **414**. Output signals can then be generated from connections **414** via outputs **416** and **418**. A connection network can therefore be configured from the plurality of connections **414**. Such a connection network is generally associated with one or more neuron-like nodes.

The connection network also comprises a plurality of interconnected nanoconnections, wherein each nanoconnection thereof is strengthened or weakened according to an application of an electric field. A connection network is not possible if built in one layer because the presence of one connection can alter the electric field so that other connections between adjacent electrodes could not be formed. Instead, such a connection network can be built in layers, so that each connection thereof can be formed without being influenced by field disturbances resulting from other connections. This can be seen in FIG. **5**.

FIG. **5** illustrates a schematic diagram of a physical neural network **500** that can be created without disturbances, in accordance with a preferred embodiment. Physical neural network **500** is composed of a first layer **558** and a second layer **560**. A plurality of inputs **502**, **504**, **506**, **508**, and **510** are respectively provided to layers **558** and **560** respectively via a plurality of input lines **512**, **514**, **516**, **518**, and **520** and a plurality of input lines **522**, **524**, **526**, **528**, and **530**. Input lines **512**, **514**, **516**, **518**, and **520** are further coupled to input lines **532**, **534**, **536**, **538**, and **540** such that each line **532**, **534**, **536**, **538**, and **540** is respectively coupled to nanoconnections **572**, **574**, **576**, **578**, and **580**. Thus, input line **532** is connected to nanconnections **572**. Input line **534** is connected to nanoconnections **574**, and input line **536** is connected to nanoconnections **576**. Similarly, input line **538** is connected to nanconnections **578**, and input line **540** is connected to nanoconnections **580**.

Nanconnections **572**, **574**, **576**, **578**, and **580** may comprise nanoconductors such as, for example, nanotubes and/or nanowires. Nanoconnections **572**, **574**, **576**, **578**, and **580** thus comprise one or more nanoconductors. Additionally, input lines **522**, **524**, **526**, **528**, and **530** are respectively coupled to a plurality of input lines **542**, **544**, **546**, **548** and **550**, which are in turn each respectively coupled to nanoconnections **582**, **584**, **586**, **588**, and **590**. Thus, for example, input line **542** is connected to nanoconnections **582**, while input line **544** is connected to nanoconnections **584**. Similarly, input line **546** is connected to nanoconnections **586** and input line **548** is connected to nanoconnections **588**. Additionally, input line **550** is connected to nanconnections **590**. Box **556** and **554** generally represent simply the output and are thus illustrated connected to outputs **562** and **568**. In other words, outputs **556** and **554** respectively comprise outputs **562** and **568**. The aforementioned input lines and associated components thereof actually comprise physical electronic components, including conducting input and output lines and physical nanoconnections, such as nanotubes and/or nanowires.

Thus, the number of layers **558** and **560** equals the number of desired outputs **562** and **568** from physical neural network **500**. In the previous two figures, every input was potentially connected to every output, but many other configurations are possible. The connection network can be made of any electrically conducting material, although the physics of it requires that they be very small so that they will align with a practical voltage. Carbon nanotubes or any conductive nanowire can be implemented in accordance with the physical neural network described herein. Such

components can form connections between electrodes by the presence of an electric field. For example, the orientation and purification of carbon nanotubes has been demonstrated using ac electrophoresis in isopropyl alcohol, as indicated in "Orientation and purification of carbon nanotubes using ac electrophoresis" by Yamamoto et al., J. Phys. D: Applied Physics, 31 (1998), 34–36. Additionally, an electric-field assisted assembly technique used to position individual nanowires suspended in an electric medium between two electrodes defined lithographically on an SiO_2 substrate is indicated in "Electric-field assisted assembly and alignment of metallic nanowires," by Smith et al., Applied Physics Letters, Vol. 77, Num. 9, Aug. 28, 2000. Such references are referred to herein for edification and illustrative purposes only.

The only general requirements for the conducting material utilized to configure the nanoconductors are that such conducting material should preferably conduct electricity, and a dipole should preferably be induced in the material when in the presence of an electric field. Alternatively, the nanoconductors utilized in association with the physical neural network described herein can be configured to include a permanent dipole that is produced by a chemical means, rather than a dipole that is induced by an electric field.

Therefore, it should be appreciated by those skilled in the art that a connection network could also be comprised of other conductive particles that may be developed or found useful in the nanotechnology arts. For example, carbon particles (or "dust") may also be used as nanoconductors in place of nanowires or nanotubes. Such particles may include bearings or granule-like particles.

A connection network can be constructed as follows: A voltage is applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture could be made of many things. The only requirements are that the conducting wires must be suspended in the solvent, either dissolved or in some sort of suspension, free to move around; the electrical conductance of the substance must be less than the electrical conductance of the suspended conducting wire; and the viscosity of the substance should not be too much so that the conducting wire cannot move when an electric field is applied.

The goal for such a connection network is to develop a network of connections of just the right values so as to satisfy the particular signal-processing requirement—exactly what a neural network does. Such a connection network can be constructed by applying a voltage across a space occupied by the mixture mentioned. To create the connection network, the input terminals are selectively raised to a positive voltage while the output terminals are selectively grounded. Thus, connections can gradually form between the inputs and outputs. The important requirement that makes the physical neural network functional as a neural network is that the longer this electric field is applied across a connection gap, or the greater the frequency or amplitude, the more nanotubes and/or nanowires and/or particles align and the stronger the connection thereof becomes. Thus, the connections that are utilized most frequently by the physical neural network become the strongest.

The connections can either be initially formed and have random resistances or no connections may be formed at all. By initially forming random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up from scratch. Depending on the rate of connection decay, having initial random connections could prove faster, although not necessarily. The connection network can adapt itself to the requirements of a given situation regardless of the initial state of the connections. Either initial condition will work, as connections that are not used will "dissolve" back into solution. The resistance of the connection can be maintained or lowered by selective activations of the connection. In other words, if the connection is not used, it will fade away, analogous to the connections between neurons in a human brain. The temperature of the solution can also be maintained at a particular value so that the rate that connections fade away can be controlled. Additionally an electric field can be applied perpendicular to the connections to weaken them, or even erase them out altogether (i.e., as in clear, zero, or reformatting of a "disk").

The nanoconnections may or may not be arranged in an orderly array pattern. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, due to the non-linear nature of neural networks, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network **600** that can be implemented in accordance with an alternative embodiment. Note that in FIGS. **5** and **6**, like parts are indicated by like reference numerals. Thus, physical neural network **600** can be configured, based on physical neural network **500** illustrated in FIG. **5**. In FIG. **6**, inputs **1**, **2**, **3**, **4**, and **5** are indicated, which are respectively analogous to inputs **502**, **504**, **506**, **508**, and **510** illustrated in FIG. **5**. Outputs **562** and **568** are provided to a plurality of electrical components to create a first output **626** (i.e., Output **1**) and a second output **628** (i.e., Output **2**). Output **562** is tied to a resistor **606**, which is labeled R**2** and a diode **616** at node A. Output **568** is tied to a resistor **610**, which is also labeled R**2** and a diode **614** at node C. Resistors **606** and **610** are each tied to a ground **602**.

Diode **616** is further coupled to a resistor **608**, which is labeled R**3**, and first output **626**. Additionally, resistor **608** is coupled to ground **602** and an input to an amplifier **618**. An output from amplifier **618**, as indicated at node B and dashed lines thereof, can be tied back to node A. A desired output **622** from amplifier **618** is coupled to amplifier **618** at node H. Diode **614** is coupled to a resistor **612** at node F. Note that resistor **612** is labeled R**3**. Node F is in turn coupled to an input of amplifier **620** and to second output **628** (i.e., Output **2**). Diode **614** is also connected to second output **628** and an input to amplifier **620** at second output **628**. Note that second output **628** is connected to the input to amplifier **620** at node F. An output from amplifier **620** is further coupled to node D, which in turn is connected to node C. A desired output **624**, which is indicated by a dashed line in FIG. **6**, is also coupled to an input of amplifier **620** at node E.

In FIG. **6**, the training of physical neural network **600** can be accomplished utilizing, for example, op-amp devices (e.g., amplifiers **618** and **620**). By comparing an output (e.g., first output **626**) of physical neural network **600** with a desired output (e.g., desired output **622**), the amplifier (e.g., amplifier **618**) can provide feedback and selectively

strengthen connections thereof. For instance, suppose it is desired to output a voltage of +V at first output **626** (i.e., Output **1**) when inputs **1** and **4** are high. When inputs **1** and **4** are taken high, also assume that first output **626** is zero. Amplifier **618** can then compare the desired output (+V) with the actual output (**0**) and output −V. In this case, −V is equivalent to ground.

The op-amp outputs and grounds the pre-diode junction (i.e., see node A) and causes a greater electric field across inputs **1** and **4** and the layer **1** output. This increased electric field (larger voltage drop) can cause the nanoconductors in the solution between the electrode junctions to align themselves, aggregate, and form a stronger connection between the 1 and 4 electrodes. Feedback can continue to be applied until output of physical neural network **600** matches the desired output. The same procedure can be applied to every output.

In accordance with the aforementioned example, assume that Output **1** was higher than the desired output (i.e., desired output **622**). If this were the case, the op-amp output can be +V and the connection between inputs **1** and **4** and layer one output can be raised to +V. Columbic repulsions between the nanoconductors can force the connection apart, thereby weakening the connection. The feedback will then continue until the desired output is obtained. This is just one training mechanism. One can see that the training mechanism does not require any computations, because it is a simple feedback mechanism.

Such a training mechanism, however, may be implemented in many different forms. Basically, the connections in a connection network must be able to change in accordance with the feedback provided. In other words, the very general notion of connections being strengthened or connections being weakened in a physical system is the essence of a physical neural network (i.e., Knowm). Thus, it can be appreciated that the training of such a physical neural network may not require a "CPU" to calculate connection values thereof. The Knowm can adapt itself. Complicated neural network solutions could be implemented very rapidly "on the fly", much like a human brain adapts as it performs.

The physical neural network disclosed herein thus has a number of broad applications. The core concept of a Knowm, however, is basic. The very basic idea that the connection values between electrode junctions by nanoconductors can be used in a neural network devise is all that required to develop an enormous number of possible configurations and applications thereof.

Another important feature of a physical neural network is the ability to form negative connections. This is an important feature that makes possible inhibitory effects useful in data processing. The basic idea is that the presence of one input can inhibit the effect of another input. In artificial neural networks as they currently exist, this is accomplished by multiplying the input by a negative connection value. Unfortunately, with a Knowm-based device, the connection may only take on zero or positive values under such a scenario.

In other words, either there can be a connection or no connection. A connection can simulate a negative connection by dedicating a particular connection to be negative, but one connection cannot begin positive and through a learning process change to a negative connection. In general, if starts positive, it can only go to zero. In essence, it is the idea of possessing a negative connection initially that results in the simulation, because this does not occur in a human brain. Only one type of signal travels through axon/dendrites in a human brain. That signal is transferred into the flow of a neurotransmitter whose effect on the postsynaptic neuron can be either excitatory or inhibitory, depending on the neuron.

One method for solving this problem is to utilize two sets of connections for the same output, having one set represent the positive connections and the other set represent the negative connections. The output of these two layers can be compared, and the layer with the greater output will output either a high signal or a low signal, depending on the type of connection set (inhibitory or excitatory). This can be seen in FIG. **7**.

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network **700** that can be implemented in accordance with an alternative embodiment. Physical neural network **700** thus comprises a plurality of inputs **702** (not necessarily binary) which are respectively fed to layers **704**, **706**, **708**, and **710**. Each layer is analogous to the layers depicted earlier, such as for example layers **558** and **560** of FIG. **5**. An output **713** of layer **704** can be connected to a resistor **712**, a transistor **720** and a first input **727** of amplifier **726**. Transistor **720** is generally coupled between ground **701** and first input **727** of amplifier **726**. Resistor **712** is connected to a ground **701**. Note that ground **701** is analogous to ground **602** illustrated in FIG. **6** and ground **210** depicted in FIG. **2**. A second input **729** of amplifier **726** can be connected to a threshold voltage **756**. The output of amplifier **726** can in turn be fed to an inverting amplifier **736**.

The output of inverting amplifier **736** can then be input to a NOR device **740**. Similarly, an output **716** of layer **706** may be connected to resistor **714**, transistor **733** and a first input **733** of an amplifier **728**. A threshold voltage **760** is connected to a second input **737** of amplifier **728**. Resistor **714** is generally coupled between ground **701** and first input **733** of amplifier **728**. Note that first input **733** of amplifier **728** is also generally connected to an output **715** of layer **706**. The output of amplifier **728** can in turn be provided to NOR device **740**. The output from NOR device **740** is generally connected to a first input **745** of an amplifier **744**. An actual output **750** can be taken from first input **745** to amplifier **744**. A desired output **748** can be taken from a second input **747** to amplifier **744**. The output from amplifier **744** is generally provided at node A, which in turn is connected to the input to transistor **720** and the input to transistor **724**. Note that transistor **724** is generally coupled between ground **701** and first input **733** of amplifier **728**. The second input **731** of amplifier **728** can produce a threshold voltage **760**.

Layer **708** provides an output **717** that can be connected to resistor **716**, transistor **725** and a first input **737** to an amplifier **732**. Resistor **716** is generally coupled between ground **701** and the output **717** of layer **708**. The first input **737** of amplifier **732** is also electrically connected to the output **717** of layer **708**. A second input **735** to amplifier **732** may be tied to a threshold voltage **758**. The output from amplifier **732** can in turn be fed to an inverting amplifier **738**. The output from inverting amplifier **738** may in turn be provided to a NOR device **742**. Similarly, an output **718** from layer **710** can be connected to a resistor **719**, a transistor **728** and a first input **739** of an amplifier **734**. Note that resistor **719** is generally coupled between node **701** and the output **719** of layer **710**. A second input **741** of amplifier **734** may be coupled to a threshold voltage **762**. The output from of NOR device **742** is generally connected to a first input **749** of an amplifier **746**. A desired output **752** can be taken from a second input **751** of amplifier **746**. An actual output **754** can be taken from first input **749** of amplifier **746**.

The output of amplifier **746** may be provided at node B, which in turn can be tied back to the respective inputs to transistors **725** and **728**. Note that transistor **725** is generally coupled between ground **701** and the first input **737** of amplifier **732**. Similarly, transistor **728** is generally connected between ground **701** and the first input **739** of amplifier **734**.

Note that transistors **720**, **724**, **725** and/or **728** each can essentially function as a switch to ground. A transistor such as, for example, transistor **720**, **724**, **725** and/or **728** may comprise a field-effect transistor (FET) or another type of transistor, such as, for example, a single-electron transistor (SET). Single-electron transistor (SET) circuits are essential for hybrid circuits combining quantum SET devices with conventional electronic devices. Thus, SET devices and circuits may be adapted for use with the physical neural network of the embodiments. This is particularly important because as circuit design rules begin to move into regions of the sub-100 nanometer scale, where circuit paths are only 0.001 of the thickness of a human hair, prior art device technologies will begin to fail, and current leakage in traditional transistors will become a problem. SET offers a solution at the quantum level, through the precise control of a small number of individual electrons. Transistors such as transistors **720**, **724**, **725** and/or **728** can also be implemented as carbon nanotube transistors.

A truth table for the output of circuit **700** is illustrated at block **780** in FIG. **7**. As indicated at block **780**, when an excitatory output is high and the inhibitory output is also high, the final output is low. When the excitatory output is high and the inhibitory output is low, the final output is high. Similarly, when the excitatory output is low and the inhibitory output is high, the final output is low. When the excitatory output is low and the inhibitory output is also low, the final output is low. Note that layers **704** and **708** may thus comprise excitatory connections, while layers **706** and **710** may comprise inhibitory connections.

For every desired output, two sets of connections are used. The output of a two-diode neuron can be fed into an op-amp (e.g., a comparator). If the output that the op-amp receives is low when it should be high, the op-amp outputs a low signal. This low signal can cause the transistors (e.g., transistors **720**, **725**) to saturate and ground out the pre-diode junction for the excitatory diode. Such a scenario can cause, as indicated previously, an increase in the voltage drop across those connections that need to increase their strength. Note that only those connections going to the excitatory diode are strengthened. Likewise, if the desired output were low when the actual output was high, the op-amp can output a high signal. This can cause the inhibitory transistor (e.g., an NPN transistor) to saturate and ground out the neuron junction of the inhibitory connections. Those connections going to the inhibitory diode can thereafter strengthen.

At all times during the learning process, a weak alternating electric field can be applied perpendicular to the connections. This can cause the connections to weaken by rotating the nanotube perpendicular to the connection direction. This perpendicular field is important because it can allow for a much higher degree of adaptation. To understand this, one must realize that the connections cannot (practically) keep getting stronger and stronger. By weakening those connections not contributing much to the desired output, we decrease the necessary strength of the needed connections and allow for more flexibility in continuous training. This perpendicular alternating voltage can be realized by the addition of two electrodes on the outer extremity of the connection set, such as plates sandwiching the connections (i.e., above and below). Other mechanisms, such as increasing the temperature of the nanotube suspension could also be used for such a purpose, although this method is perhaps a little less controllable or practical.

The circuit depicted in FIG. **7** can be separated into two separate circuits. The first part of the circuit can be composed of nanotube connections, while the second part of the circuit comprises the "neurons" and the learning mechanism (i.e., op-amps/comparator). The learning mechanism on first glance appears similar to a relatively standard circuit that could be implemented on silicon with current technology. Such a silicon implementation can thus comprise the "neuron" chip. The second part of the circuit (i.e., the connections) is thus a new type of chip, although it could be constructed with current technology. The connection chip can be composed of an orderly array of electrodes spaced anywhere from, for example, 100 nm to 1 μm or perhaps even further. In a biological system, one talks of synapses connecting neurons. It is in the synapses where the information is processed, (i.e., the "connection weights"). Similarly, such a chip can contain all of the synapses for the physical neural network. A possible arrangement thereof can be seen in FIG. **8**.

FIG. **8** illustrates a schematic diagram of a chip layout **800** for a connection network that may be implemented in accordance with an alternative embodiment. FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., connection network **800**) that can be implemented in accordance with one or more embodiments. Chip layout **800** includes an input array composed of plurality of inputs **801**, **802**, **803**, **804**, and **805**, which are provided to a plurality of layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. A plurality of outputs **802** can be derived from layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. Inputs **801** can be coupled to layers **806** and **807**, while inputs **802** can be connected to layers **808** and **809**. Similarly, inputs **803** can be connected to layers **810** and **811**. Also, inputs **804** can be connected to layers **812** and **813**. Inputs **805** are generally connected to layers **814** and **815**.

Similarly, such an input array can includes a plurality of inputs **831**, **832**, **833**, **834** and **835** which are respectively input to a plurality of layers **816**, **817**, **818**, **819**, **820**, **821**, **822**, **823**, **824** and **825**. Thus, inputs **831** can be connected to layers **816** and **817**, while inputs **832** are generally coupled to layers **818** and **819**. Additionally, inputs **833** can be connected to layers **820** and **821**. Inputs **834** can be connected to layers **822** and **823**. Finally, inputs **835** are connected to layers **824** and **825**. Arrows **828** and **830** represent a continuation of the aforementioned connection network pattern. Those skilled in the art can appreciate, of course, that chip layout **800** is not intended to represent an exhaustive chip layout or to limit the scope of the invention. Many modifications and variations to chip layout **800** are possible in light of the teachings herein without departing from the scope of the embodiments. It is contemplated that the use of a chip layout, such as chip layout **800**, can involve a variety of components having different characteristics.

Preliminary calculations based on a maximum etching capability of 200 nm resolution indicated that over 4 million synapses could fit on an area of approximately 1 cm^2. The smallest width that an electrode can possess is generally based on current lithography. Such a width may of course change as the lithographic arts advance. This value is actually about 70 nm for state-of-the-art techniques currently. These calculations are of course extremely conservative, and are not considered a limiting feature of the embodiments. Such calculations are based on an electrode

with, separation, and gap of approximately 200 nm. For such a calculation, for example, 166 connection networks comprising 250 inputs and 100 outputs can fit within a one square centimeter area.

If such chips are stacked vertically, an untold number of synapses could be attained. This is two to three orders of magnitude greater than some of the most capable neural network chips out there today, chips that rely on standard methods to calculate synapse weights. Of course, the geometry of the chip could take on many different forms, and it is quite possible (based on a conservative lithography and chip layout) that many more synapses could fit in the same space. The training of a chip this size would take a fraction of the time of a comparably sized traditional chip using digital technology.

The training of such a chip is primarily based on two assumptions. First, the inherent parallelism of a physical neural network (i.e., a Knowm) can permit all training sessions to occur simultaneously, no matter now large the associated connection network. Second, recent research has indicated that near perfect aligning of nanotubes can be accomplished in approximately 15 minutes. If one considers that the input data, arranged as a vector of binary "high's" and "low's" is presented to the Knowm simultaneously, and that all training vectors are presented one after the other in rapid succession (e.g., perhaps 100 MHz or more), then each connection would "see" a different frequency in direct proportion to the amount of time that its connection is required for accurate data processing (i.e., provided by a feedback mechanism). Thus, if it only takes approximately 15 minutes to attain an almost perfect state of alignment, then this amount of time would comprise the longest amount of time required to train, assuming that all of the training vectors are presented during that particular time period.

FIG. **9** illustrates a flow chart **900** of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment. Initially, as indicated at block **902**, a connection gap is created from a connection network structures. As indicated earlier, the goal for such a connection network is generally to develop a network of connections of "just" the right values to satisfy particular information processing requirements, which is precisely what a neural network accomplishes. As illustrated at block **904**, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent," and also a suspension.

The solvent utilized can comprise a volatile liquid that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution may be sandwiched between wafers of silicon or other materials. If the fluid has a melting point that is approximately at room temperature, then the viscosity of the fluid could be controlled easily. Thus, if it is desired to lock the connection values into a particular state, the associated physical neural network (i.e., Knowm) may be cooled slightly until the fluid freezes. The term "solvent" as utilized herein thus can include fluids such as for example, toluene, hexadecane, mineral oil, etc. Note that the solution in which the nanoconductors (i.e., nanoconnections) are present should generally comprise a dielectric. Thus, when the resistance between the electrodes is measured, the conductivity of the nanoconductors can be essentially measured, not that of the solvent. The nanoconductors can be suspended in the solution or can alternately lie on the bottom

surface of the connection gap. The solvent may also be provided in the form of a gas.

As illustrated thereafter at block **906**, the nanoconductors must be suspended in the solvent, either dissolved or in a suspension of sorts, but generally free to move around, either in the solution or on the bottom surface of the gap. As depicted next at block **908**, the electrical conductance of the solution must be less than the electrical conductance of the suspended nanoconductor(s). Similarly, the electrical resistance of the solution is greater than the electrical resistance of the nanoconductor.

Next, as illustrated at block **910**, the viscosity of the substance should not be too much so that the nanoconductors cannot move when an electric field (e.g., voltage) is applied. Finally, as depicted at block **912**, the resulting solution of the "solvent" and the nanoconductors is thus located within the connection gap.

Note that although a logical series of steps is illustrated in FIG. **9**, it can be appreciated that the particular flow of steps can be re-arranged. Thus, for example, the creation of the connection gap, as illustrated at block **902**, may occur after the preparation of the solution of the solvent and nanoconductor(s), as indicated at block **904**. FIG. **9** thus represents merely possible series of steps, which may be followed to create a connection network. A variety of other steps may be followed as long as the goal of achieving a connection network is achieved. Similar reasoning also applies to FIG. **10**.

FIG. **10** illustrates a flow chart **1000** of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment. As indicated at block **1002**, an electric field can be applied across the connection gap discussed above with respect to FIG. **9**. The connection gap can be occupied by the solution discussed above. As indicated thereafter at block **1004**, to create the connection network, the input terminals can be selectively raised to a positive voltage while the output terminals are selectively grounded. As illustrated thereafter at block **1006**, connections thus form between the inputs and the outputs. The important requirements that make the resulting physical neural network functional as a neural network is that the longer this electric field is applied across the connection gap, or the greater the frequency or amplitude, the more nanoconductors align and the stronger the connection becomes. Thus, the connections that get utilized the most frequently become the strongest.

As indicated at block **1008**, the connections can either be initially formed and have random resistances or no connections will be formed at all. By forming initial random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up as much. Depending on the rate of connection decay, having initial random connections could prove to be a faster method, although not necessarily. A connection network can adapt itself to whatever is required regardless of the initial state of the connections. Thus, as indicated at block **1010**, as the electric field is applied across the connection gap, the more the nonconductor(s) will align and the stronger the connection becomes. Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block **1012**. As illustrated at block **1014**, the resistance of the connection can be maintained or lowered by selective activations of the connections. In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain.

The neurons in a human brain, although seemingly simple when viewed individually, interact in a complicated network that computes with both space and time. The most basic picture of a neuron, which is usually implemented in technology, is a summing device that adds up a signal. Actually, this statement can be made even more general by stating that a neuron adds up a signal in discrete units of time. In other words, every group of signals incident upon the neuron can be viewed as occurring in one moment in time. Summation thus occurs in a spatial manner. The only difference between one signal and another signal depends on where such signals originate. Unfortunately, this type of data processing excludes a large range of dynamic, varying situations that cannot necessarily be broken up into discrete units of time.

The example of speech recognition is a case in point. Speech occurs in the time domain. A word is understood as the temporal pronunciation of various syllables. A sentence is composed of the temporal separation of varying words. Thoughts are composed of the temporal separation of varying sentences. Thus, for an individual to understand a spoken language at all, a syllable, word, sentence or thought must exert some type of influence on another syllable, word, sentence or thought. The most natural way that one sentence can exert any influence on another sentence, in the light of neural networks, is by a form of temporal summation. That is, a neuron "remembers" the signals it received in the past.

The human brain accomplishes this feat in an almost trivial manner. When a signal reaches a neuron, the neuron has an influx of ions rush through its membrane. The influx of ions contributes to an overall increase in the electrical potential of the neuron. Activation is achieved when the potential inside the cell reaches a certain threshold. The one caveat is that it takes time for the cell to pump out the ions, something that it does at a more or less constant rate. So, if another signal arrives before the neuron has time to pump out all of the ions, the second signal will add with the remnants of the first signal and achieve a raised potential greater than that which could have occurred with only the second signal. The first signal influences the second signal, which results in temporal summation.

Implementing this in a technological manner has proved difficult in the past. Any simulation would have to include a "memory" for the neuron. In a digital representation, this requires data to be stored for every neuron, and this memory would have to be accessed continually. In a computer simulation, one must discritize the incoming data, since operations (such as summations and learning) occur serially. That is, a computer can only do one thing at a time. Transformations of a signal from the time domain into the spatial domain require that time be broken up into discrete lengths, something that is not necessarily possible with real-time analog signals in which no point exists within a time-varying signal that is uninfluenced by another point.

A physical neural network, however, is generally not digital. A physical neural network is a massively parallel analog device. The fact that actual molecules (e.g., nanoconductors) must move around (in time) makes temporal summation a natural occurrence. This temporal summation is built into the nanoconnections. The easiest way to understand this is to view the multiplicity of nanoconnections as one connection with one input into a neuron-like node (Op-amp, Comparator, etc.). This can be seen in FIG. **11**.

FIG. **11** illustrates a schematic diagram of a circuit **1100** illustrating temporal summation within a neuron, in accordance with a preferred embodiment. As indicated in FIG. **11**, an input **1102** can be provided to nanoconnections **1104**, which in turn can provide a signal, which can be input to an amplifier **1110** (e.g., op amp) at node B. A resistor **1106** can be connected to node A, which in turn is electrically equivalent to node B. Node B can be connected to a negative input of amplifier **1100**. Resistor **1108** can also be connected to a ground **1108**. Amplifier **1110** provides output **1114**. Note that although nanoconnections **1104** is referred to in the plural it can be appreciated that nanoconnections **1104** can comprise a single nanoconnection or a plurality of nanoconnections. For simplicity sake, however, the plural form is used to refer to nanoconnections **1104**.

Input **1102** can be provided by another physical neural network (i.e., Knowm) to cause increased connection strength of nanoconnections **1104** over time. This input would most likely arrive in pulses, but could also be continuous. A constant or pulsed electric field perpendicular to the connections can serve to constantly erode the connections, so that only signals of a desired length or amplitude can cause a connection to form. Once the connection is formed, the voltage divider formed by nanoconnection **1104** and resistor **1106** can cause a voltage at node A in direct proportion to the strength of nanoconnections **1104**. When the voltage at node A reaches a desired threshold, the amplifier (i.e., an op-amp and/or comparator), will output a high voltage (i.e., output **1114**). The key to the temporal summation is that, just like a real neuron, it takes time for the electric field to breakdown the nanoconnections **1104**, so that signals arriving close in time will contribute to the firing of the neuron (i.e., op-amp, comparator, etc.). Temporal summation has thus been achieved. The parameters of the temporal summation could be adjusted by the amplitude and frequency of the input signals and the perpendicular electric field.

FIG. **12** illustrates a block diagram illustrating a pattern recognition system **1200**, which may be implemented with a physical neural network device **1222**, in accordance with an alternative embodiment. Note that pattern recognition system **1200** can be implemented as a speech recognition system. Although pattern recognition system **1200** is depicted herein in the context of speech recognition, a physical neural network device (i.e., a Knowm device) may be implemented with other pattern recognition systems, such as visual and/or imaging recognition systems. FIG. **12** thus does not comprise a limiting feature of the embodiments and is presented for general edification and illustrative purposes only. Those skilled in the art can appreciate that the diagram depicted in FIG. **12** may be modified as new applications and hardware are developed. The development or use of a pattern recognition system such as pattern recognition system **1200** of FIG. **12** by no means limits the scope of the physical neural network (i.e., Knowm) disclosed herein.

FIG. **12** thus illustrates in block diagram fashion, the system structure of a speech recognition device using a neural network according to an alternative embodiment. The pattern recognition system **1200** can be provided with a CPU **1211** for performing the functions of inputting vector rows and instructor signals (vector rows) to an output layer for the learning process of a physical neural network device **1222**, and changing connection weights between respective neuron devices based on the learning process. Pattern recognition system **1200** can be implemented within the context of a data-processing system, such as, for example, a personal computer or personal digital assistant (PDA), both of which are well known in the art.

The CPU **1211** can perform various processing and controlling functions, such as pattern recognition, including but not limited to speech and/or visual recognition based on the output signals from the physical neural network device

1222. The CPU 1211 is connected to a read-only memory (ROM) 1213, a random-access memory (RAM) 1214, a communication control unit 1215, a printer 1216, a display unit 1217, a keyboard 1218, an FFT (fast Fourier transform) unit 1221, a physical neural network device 1222 and a graphic reading unit 1224 through a bus line 1220 such as a data bus line. The bus line 1220 may comprise, for example, an ISA, EISA, or PCI bus.

The ROM 1213 is a read-only memory storing various programs or data used by the CPU 1211 for performing processing or controlling the learning process, and speech recognition of the physical neural network device 1222. The ROM 1213 may store programs for carrying out the learning process according to error back-propagation for the physical neural network device or code rows concerning, for example, 80 kinds of phonemes for performing speech recognition. The code rows concerning the phonemes can be utilized as second instructor signals and for recognizing phonemes from output signals of the neuron device network. Also, the ROM 1213 can store programs of a transformation system for recognizing speech from recognized phonemes and transforming the recognized speech into a writing (i.e., written form) represented by characters.

A predetermined program stored in the ROM 1213 can be downloaded and stored in the RAM 1214. RAM 1214 generally functions as a random access memory used as a working memory of the CPU 1211. In the RAM 1214, a vector row storing area can be provided for temporarily storing a power obtained at each point in time for each frequency of the speech signal analyzed by the FFT unit 1221. A value of the power for each frequency serves as a vector row input to a first input portion of the physical neural network device 1222. Further, in the case where characters or graphics are recognized in the physical neural network device, the image data read by the graphic reading unit 1224 are stored in the RAM 1214.

The communication control unit 1215 transmits and/or receives various data such as recognized speech data to and/or from another communication control unit through a communication network 1202 such as a telephone line network, an ISDN line, a LAN, or a personal computer communication network. Network 1202 may also comprise, for example, a telecommunications network, such as a wireless communications network. Communication hardware methods and systems thereof are well known in the art.

The printer 1216 can be provided with a laser printer, a bubble-type printer, a dot matrix printer, or the like, and prints contents of input data or the recognized speech. The display unit 1217 includes an image display portion such as a CRT display or a liquid crystal display, and a display control portion. The display unit 1217 can display the contents of the input data or the recognized speech as well as a direction of an operation required for speech recognition utilizing a graphical user interface (GUI).

The keyboard 1218 generally functions as an input unit for varying operating parameters or inputting setting conditions of the FFT unit 1221, or for inputting sentences. The keyboard 1218 is generally provided with a ten-key numeric pad for inputting numerical figures, character keys for inputting characters, and function keys for performing various functions. A mouse 1219 can be connected to the keyboard 1218 and serves as a pointing device.

A speech input unit 1223, such as a microphone can be connected to the FFT unit 1221. The FFT unit 1221 transforms analog speech data input from the voice input unit 1223 into digital data and carries out spectral analysis of the digital data by discrete Fourier transformation. By performing a spectral analysis using the FFT unit 1221, the vector row based on the powers of the respective frequencies are output at predetermined intervals of time. The FFT unit 1221 performs an analysis of time-series vector rows, which represent characteristics of the inputted speech. The vector rows output by the FFT 1221 are stored in the vector row storing area in the RAM 1214.

The graphic reading unit 224, provided with devices such as a CCD (Charged Coupled Device), can be used for reading images such as characters or graphics recorded on paper or the like. The image data read by the image-reading unit 1224 are stored in the RAM 1214. Note that an example of a pattern recognition apparatus, which may be modified for use with the physical neural network described herein, is disclosed in U.S. Pat. No. 6,026,358 to Tomabechi, Feb. 16, 2000, "Neural Network, A Method of Learning of a Neural Network and Phoneme Recognition Apparatus Utilizing a Neural Network." U.S. Pat. No. 6,026,358 is incorporated herein by reference. It can be appreciated that the Tomabechi reference does not teach, suggest or anticipate the embodiments, but is discussed herein for general illustrative, background and general edification purposes only.

The implications of a physical neural network are tremendous. With existing lithography technology, many electrodes in an array such as depicted in FIG. 5 can be etched onto a wafer of silicon. The neuron-diodes, as well as the training circuitry illustrated in FIG. 6, could be built onto the same silicon wafer, although it may be desirable to have the connections on a separate chip due to the liquid solution of nanoconductors. A solution of suspended nanoconductors could be placed between the electrode connections and the chip could be packaged. The resulting "chip" would look much like a current Integrated Chip (IC) or VLSI (very large scale integrated) chips. One could also place a rather large network parallel with a computer processor as part of a larger system. Such a network, or group of networks, could add significant computational capabilities to standard computers and associated interfaces.

For example, such a chip may be constructed utilizing a standard computer processor in parallel with a large physical neural network or group of physical neural networks. A program can then be written such that the standard computer teaches the neural network to read, or create an association between words, which is precisely the same sort of task in which neural networks can be implemented. Once the physical neural network is able to read, it can be taught for example to "surf" the Internet and find material of any particular nature. A search engine can then be developed that does not search the Internet by "keywords", but instead by meaning. This idea of an intelligent search engine has already been proposed for standard neural networks, but until now has been impractical because the network required was too big for a standard computer to simulate. The use of a physical neural network (i.e., physical neural network) as disclosed herein now makes a truly intelligent search engine possible.

A physical neural network can be utilized in other applications, such as, for example, speech recognition and synthesis, visual and image identification, management of distributed systems, self-driving cars, filtering, etc. Such applications have to some extent already been accomplished with standard neural networks, but are generally limited in expense, practicality and not very adaptable once implemented. The use of a physical neural network can permit such applications to become more powerful and adaptable. Indeed, anything that requires a bit more "intelligence" could incorporate a physical neural network. One of the

primary advantages of a physical neural network is that such a device and applications thereof can be very inexpensive to manufacture, even with present technology. The lithographic techniques required for fabricating the electrodes and channels therebetween has already been perfected and implemented in industry.

Most problems in which a neural network solution is implemented are complex adaptive problems, which change in time. An example is weather prediction. The usefulness of a physical neural network is that it could handle the enormous network needed for such computations and adapt itself in real-time. An example wherein a physical neural network (i.e., Knowm) can be particularly useful is the Personal Digital Assistant (PDA). PDA's are well known in the art. A physical neural network applied to a PDA device can be advantageous because the physical neural network can ideally function with a large network that could constantly adapt itself to the individual user without devouring too much computational time from the PDA. A physical neural network could also be implemented in many industrial applications, such as developing a real-time systems control to the manufacture of various components. This systems control can be adaptable and totally tailored to the particular application, as necessarily it must.

It will be appreciated that variations of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also that various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

What is claimed is:

1. A pattern recognition system, comprising:

a neural network based on nanotechnology, wherein said neural network comprises a plurality of nanoconnections disposed within a solution within a connection network of said plurality of nanoconnections; and

a pattern input unit, which communicates with said neural network, wherein said neural network processes data input via said pattern input unit in order to recognize data patterns.

2. The system of claim **1** further comprising a processor that communicates with said neural network.

3. The system of claim **1** further comprising at least one memory unit for storing data processed by said neural network, wherein said at least one memory unit is associated with said pattern input unit.

4. The system of claim **1** wherein said at least one memory unit comprises Read Only Memory (ROM).

5. The system of claim **1** wherein said at least one memory unit comprises Random Access Memory (RAM).

6. The system of claim **1** wherein said pattern input unit comprise a speech input unit.

7. The system of claim **6** further comprising a Fast Fourier Transform (FFT) unit connected to said speech input unit, wherein said FFT unit communicates data to and from said neural network.

8. The system of claim **1** further comprising a Personal Digital Assistant (PDA) in which said neural network and said pattern input unit are located.

9. The system of claim **1** wherein said pattern input unit comprises a visual pattern input unit that interacts with said neural network to recognize visual data patterns thereof.

10. The system of claim **1** wherein said pattern input unit comprises an imaging input that interacts with said neural network to recognize imaging data patterns.

11. A pattern recognition system, comprising:

a neural network based on nanotechnology, wherein said neural network comprises a plurality of nanoconnections disposed within a solution within a connection network of said plurality of nanoconnections;

a processor that communicates with said neural network;

a pattern input unit, which communicates with said neural network, wherein said neural network processes data input via said pattern input unit in order to recognize data patterns; and

at least one memory unit for storing data processed by said neural network, wherein said at least one memory unit is associated with said pattern input unit.

12. The system of claim **11** wherein said pattern input unit comprise a speech input unit.

13. The system of claim **12** further comprising a Fast Fourier Transform (FFT) unit connected to said speech input unit, wherein said FFT unit communicates data to and from said neural network.

14. The system of claim **11** wherein said pattern input unit comprises a visual pattern input unit that interacts with said neural network to recognize visual data patterns.

15. The system of claim **11** wherein said pattern input unit comprises an imaging input that interacts with said neural network to recognize imaging data patterns.

16. A pattern recognition system, comprising:

a neural network based on utilizing nanotechnology, wherein said neural network comprises a plurality of nanoconnections disposed within a solution within a connection network of said plurality of nanoconnections;

a processor that communicates with said neural network;

a pattern input unit, which communicates with said neural network, wherein said neural network processes data input via said pattern input unit in order to recognize data patterns;

at least one memory unit for storing data processed by said neural network, wherein said at least one memory unit is associated with said pattern input unit; and

a Personal Digital Assistant (PDA) in which said neural network, said processor, said pattern input unit, and said at least one memory unit are embodied.

17. The system of claim **16** wherein said pattern input unit comprise a speech input unit.

18. The system of claim **17** further comprising a Fast Fourier Transform (FFT) unit connected to said speech input unit, wherein said FFT unit communicates data to and from said neural network.

19. The system of claim **16** wherein said pattern input unit comprises a visual pattern input unit that interacts with said neural network to recognize visual data patterns.

20. The system of claim **16** wherein said pattern input unit comprises an imaging input that interacts with said neural network to recognize imaging data patterns.

* * * * *

US007039619B2

(12) **United States Patent**
Nugent

(10) **Patent No.: US 7,039,619 B2**
(45) **Date of Patent: May 2, 2006**

(54) **UTILIZED NANOTECHNOLOGY APPARATUS USING A NEUTRAL NETWORK, A SOLUTION AND A CONNECTION GAP**

(75) Inventor: **Alex Nugent**, Santa Fe, NM (US)

(73) Assignee: **Knowm Tech, LLC**, Santa Fe, NM (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 74 days.

(21) Appl. No.: **11/045,538**

(22) Filed: **Jan. 31, 2005**

(65) **Prior Publication Data**

US 2005/0256816 A1 Nov. 17, 2005

Related U.S. Application Data

(63) Continuation of application No. 10/095,273, filed on Mar. 12, 2002, now Pat. No. 6,889,216.

(51) **Int. Cl.**
G06F 15/18 (2006.01)

(52) **U.S. Cl.** **706/15**; 706/26; 706/27

(58) **Field of Classification Search** 706/15, 706/26, 27
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

2,707,223 A	4/1955	Hollman	338/32 R
3,833,894 A	9/1974	Aviram et al.	365/151
4,802,951 A	2/1989	Clark et al.	156/630
4,926,064 A	5/1990	Tapang	706/26
4,974,146 A	11/1990	Works et al.	364/200
4,988,891 A	1/1991	Mashiko	307/201
5,315,162 A	5/1994	McHardy et al.	307/201
5,422,983 A	6/1995	Castelaz et al.	395/24
5,475,794 A	12/1995	Mashiko	395/24
5,589,692 A	12/1996	Reed	257/23
5,649,063 A	7/1997	Bose	395/22
5,670,818 A	9/1997	Forouhi et al.	257/530
5,706,404 A	1/1998	Colak	395/24
5,717,832 A	2/1998	Steimle et al.	395/24
5,761,115 A	6/1998	Kozicki et al.	365/182
5,783,840 A	7/1998	Randall et al.	257/24
5,812,993 A	9/1998	Ginosar et al.	706/26
5,896,312 A	4/1999	Kozicki et al.	365/153

(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 022 764 A1 1/2000

(Continued)

OTHER PUBLICATIONS

Peter Weiss, "Circuitry in a Nonowire: Novel Growth Method May Transform Chips," Science News Online, vol. 161, No. 6; Feb. 9, 2002.

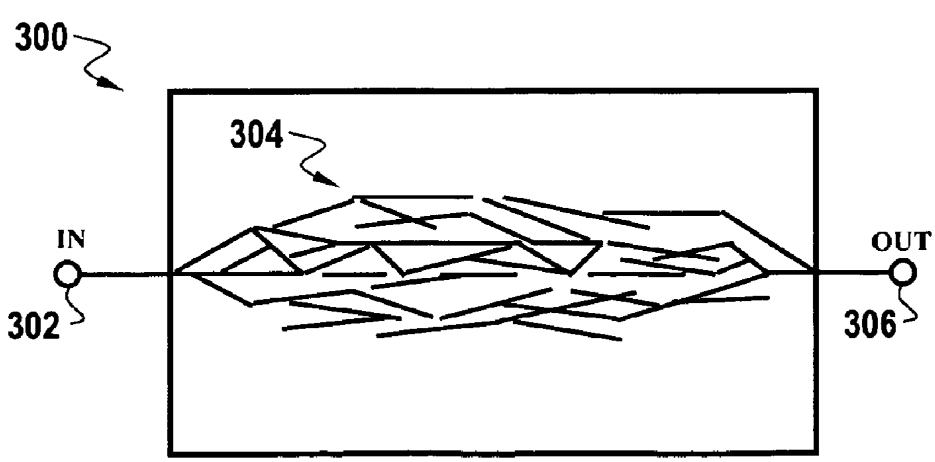
(Continued)

Primary Examiner—George Davis
(74) *Attorney, Agent, or Firm*—Kermit D. Lopez; Luis M. Ortiz; Ortiz & Lopez, PLLC

(57) **ABSTRACT**

An apparatus for maintaining components in neural network formed utilizing nanotechnology is described herein. A connection gap can be formed between two terminals. A solution comprising a melting point at approximately room temperature can be provided, wherein the solution is maintained in the connection gap and comprises a plurality of nanoparticles forming nanoconnections thereof having connection strengths thereof, wherein the solution and the connection gap are adapted for use with a neural network formed utilizing nanotechnology, such when power is removed from the neural network, the solution freezes, thereby locking into place the connection strengths.

15 Claims, 6 Drawing Sheets



U.S. PATENT DOCUMENTS

5,904,545	A	5/1999	Smith et al.	438/455
5,914,893	A	6/1999	Kozicki et al.	365/107
5,951,881	A	9/1999	Rogers et al.	216/41
5,978,782	A	11/1999	Neely	706/16
6,026,358	A	2/2000	Tomabechi	704/232
6,084,796	A	7/2000	Kozicki et al.	365/153
6,128,214	A	10/2000	Kuekes et al.	365/151
6,245,630	B1	6/2001	Ishikawa	438/393
6,248,529	B1	6/2001	Connolly	435/6
6,256,767	B1	7/2001	Kuekes et al.	716/9
6,282,530	B1	8/2001	Huang	706/41
6,294,450	B1	9/2001	Chen et al.	438/597
6,314,019	B1	11/2001	Kuekes et al.	365/151
6,330,553	B1	12/2001	Uchikawa et al.	706/2
6,335,291	B1	1/2002	Freeman	438/706
6,339,227	B1	1/2002	Ellenbogen	257/40
6,359,288	B1	3/2002	Ying et al.	257/14
6,363,369	B1	3/2002	Liaw et al.	706/15
6,383,923	B1	5/2002	Brown et al.	438/666
6,389,404	B1	5/2002	Carson et al.	706/18
6,407,443	B1	6/2002	Chen et al.	257/616
6,418,423	B1	7/2002	Kambhatla et al.	706/15
6,420,092	B1	7/2002	Yang et al.	430/311
6,422,450	B1	7/2002	Zhou et al.	228/121.85
6,423,583	B1	7/2002	Avouris et al.	438/132
6,424,961	B1	7/2002	Ayala	706/25
6,426,134	B1	7/2002	Lavin et al.	428/300.1
6,620,346	B1	9/2003	Schulz et al.	252/519.51
6,798,692	B1	9/2004	Kozicki et al.	365/174
6,855,329	B1 *	2/2005	Shakesheff et al.	424/409
2001/0004471	A1	6/2001	Zhang	427/372.2
2001/0023986	A1	9/2001	Mancevski	257/741
2001/0024633	A1	9/2001	Lee et al.	423/447.3
2001/0031900	A1	10/2001	Margrave et al.	570/126
2001/0041160	A1	11/2001	Margrave et al.	423/460
2001/0044114	A1	11/2001	Connolly	435/6
2002/0001905	A1	1/2002	Choi et al.	438/268
2002/0004028	A1	1/2002	Margrave et al.	423/447.3
2002/0004136	A1	1/2002	Gao et al.	428/367
2002/0030205	A1	3/2002	Varshavsky	257/208
2002/0075126	A1	6/2002	Reltz et al.	338/21
2002/0086124	A1	7/2002	Margrave et al.	428/36.9
2002/0090468	A1	7/2002	Goto	427/580
2002/0102353	A1	8/2002	Mauthner et al.	427/255.28
2003/0031438	A1	2/2003	Kambe et al.	385/122
2003/0177450	A1	9/2003	Nugent	716/1
2003/0236760	A1	12/2003	Nugent	706/26
2004/0039717	A1	2/2004	Nugent	706/27
2004/0150010	A1	8/2004	Snider	257/209
2004/0153426	A1	8/2004	Nugent	706/25
2004/0162796	A1	8/2004	Nugent	709/27
2004/0193558	A1	9/2004	Nugent	706/25

FOREIGN PATENT DOCUMENTS

EP	1 046 613	A2	4/2000
EP	1 100 106	A2	5/2001
EP	1 069 206	A2	7/2001
EP	1 115 135	A1	7/2001
EP	1 134 304	A2	9/2001
RU	2071126	C1	6/1996
WO	WO 00/44094		7/2000
WO	WO 03/017282	A1	8/2001
WO	WO 30/017282	A1 *	8/2001

OTHER PUBLICATIONS

Press Release, "Nanowire-based electronics and optics comes one step closer," Eureka Alert, American Chemical Society; Feb. 1, 2002.

Weeks et al., "High-pressure nanolithography using low-energy electrons from a scanning tunneling microscope," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 38-42; Dec. 12, 2001.

CMP Cientifica, "Nanotech: the tiny revolution"; CMP Cientifica, Nov. 2001.

Diehl, et al., "Self-Assembled, Deterministic Carbon Nanotube Wiring Networks," Angew. Chem. Int. Ed. 2002, 41., No. 2; Received Oct. 22, 2001.

G. Pirio, et al., "Fabrication and electrical characteristics of carbon nanotube field emission microcathodes with an integrated gate electrode," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 1-4, Oct. 2, 2001.

Leslie Smith, "An Introduction to Neural Networks," Center for Cognitive and Computational Neuroscience, Dept. of Computing & Mathematics, University of Stirling, Oct. 25, 1996; http//www.cs.stir.ac.uk/~Iss/NNIntro/InvSlides.html.

V. Derycke et al., "Carbon Nanotube Inter-and Intramolecular Logic Gates," American Chemical Society, Nano Letters, XXXX, vol. 0, No. 0, A-D.

Mark K. Anderson, "Mega Steps Toward the Nanochip," Wired News, Apr. 27, 2001.

Collins et al., "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," Science, vol. 292, pp. 706-709, Apr. 27, 2001.

Landman et al., "Metal-Semiconductor Nanocontacts: Silicon Nanowires," Physical Review Letters, vol. 85, No. 9, Aug. 28, 2000.

John G. Spooner, "Tiny tubes mean big chip advances," Cnet News.com Tech News First, Apr. 26, 2001.

Jeong-Mi Moon et al., "High-Yield Purification Process of Singlewalled Carbon Nanotubes," J. Phys. Chem. B 2001, 105, pp. 5677-5681.

"A New Class of Nanostructure: Semiconducting Nanobelts Offer Potential for Nanosensors and Nanoelectronics," Mar. 12, 2001, http://www.sciencedaily.com/releases/2001/03/010309080953.htm.

Hermanson et al., "Dielectrophoretic Assembly of Electrically Functional Microwires from Nanoparticle Suspensions," Materials Science, vol. 294, No. 5544, Issue of Nov. 2, 2001, pp. 1082-1086.

Press Release, "Toshiba Demonstrates Operation of Single-Electron Transistor Circuit at Room Temperature," Toshiba, Jan. 10, 2001.

J. Appenzeller et al., "Optimized contact configuration for the study of transport phenomena in ropes of single-wall carbon nanotubes," Applied Physics Letters, vol. 78, No. 21, pp. 3313-3315, May 21, 2001.

David Rotman, "Molecular Memory, Replacing silicon with organic molecules could mean tiny supercomputers," Technology Review, May 2001, p. 46.

Westervelt et al., "Molecular Electronics," NSF Functional Nanostructures Grant 9871810, NSF Partnership in Nanotechnology Conference, Jan. 29-30, 2001; http://www.unix.oit.umass.edu/~nano/NewFiles/FN19_Harvard.pdf.

Niyogi et al., "Chromatographic Purification of Soluble Single-Walled Carbon Nanotubes (s-SWNTs)," J. Am. Chem. Soc 2001, 123, pp. 733-734, Received Jul. 10, 2000.

Duan et al., "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices," Nature, vol. 409, Jan. 4, 2001, pp. 67-69.

Paulson, et al., "Tunable Resistance of a Carbon Nanotube-Graphite Interface," Science, vol. 290, Dec. 1, 2000, pp. 1742-1744.

Wei et al., "Reliability and current carrying capacity of carbon nanotubes," Applied Physics Letters, vol. 79, No. 8, Aug. 20, 2001, pp. 1172-1174.

Collins et al., "Nanotubes for Electronics," Scientific American, Dec. 2000, pp. 62-69.
Avouris et al., "Carbon nanotubes: nanomechanics, manipulation, and electronic devices," Applied Surface Science 141 (1999), pp. 201-209.
Smith et al., "Electric-field assisted assembly and alignment of metallic nanowires," Applied Physics Letters, vol. 77, No. 9, Aug. 28, 2000, pp. 1399-1401.
Hone et al., "Electrical and thermal transport properties of magnetically aligned single wall carbon nanotube filmm," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 666-668.
Smith et al., "structural anisotropy of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 663-665.
Andriotis et al., "Various bonding configurations of transition-metal atoms on carbon nanotubes: Their effect on contact resistance," Applied Physics Letters, vol. 76, No. 26, jun. 26, 2000, pp. 3890-3892.
Chen et al., "Aligning single-wall carbon nanotubes with an alternating-current electric field," Applied Physics Letters, vol. 78, No. 23, Jun. 4, 2001, pp 3714-3716.
Bezryadin et al., "Self-assembled chains of graphitized carbon nanoparticles," Applied Physics Letters, vol. 73, No. 18, May 3, 1999, pp. 2699-2701.
Bezryadin et al., "Evolution of avalanche conducting states in electrorheological liquids," Physical Review E, vol. 59, No. 6, Jun. 1999. pp. 6896-6901.
Liu et al., "Fullerene Pipes," Science, vol. 280, May 22, 1998, pp. 1253-1255.
Yamamoto et al., "Orientation and purification of carbon nanotubes using ac electrophoresis," J. Phys. D: Appl. Phys 31 (1998) L34-L36.
Bandow et al., "Purification of Single-Wall Carbon Nanotubes by Microfiltration," J. Phys. Chem. B 1997, 101, pp. 8839-8842.
Tohji et al., "Purifying single walled nanotubes," Nature, vol. 383, Oct. 24, 1996, p. 679.
Dejan Rakovic, "Hierarchical Neural Networks and Brainwaves: Towards a Theory of Consciousness," Brain & Consciousness: Proc. ECPD Workship (ECPD, Belgrade, 1997), pp. 189-204.
Dave Anderson & George McNeill, "Artificial Neural Networks technology," A DACS (Data & Analysis Center for Software) State-of-the-Art Report, Contract No. F30602-89-C-0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, Aug. 20, 1992.
Greg Mitchell, "Sub-50 nm Device Fabrication Strategies," Project No. 890-00, Cornell Nanofabrication Facility, Electronics—p. 90-91, National Nanofabrication Users Network.
John-William DeClaris, "An Introduction to Neural Networks," http://www.ee.umd.edu/medlab/neural/nn1.html.
"Neural Networks," StatSoft Inc., http://www.statsoftinc.com/textbook/stevnet.html.
Stephen Jones, "Neural Networks and the Computation Brain or Maters relating to Artificial Intelligence," The Brain Project, http://www.culture.com.au/brain_proj/neur_net.htm.
David W. Clark, "An Introduction to Neural Networks"; http://members.home.net/neuralnet/introtonn/index.htm.
"A Basic Introduction to Neural Networks"; http://blizzard.gis.uiuc.edu/htmldocs/Neural/neural.html.
Meyer et al., "Computational neural networks: a general purpose tool for nanotechnology," Abstract, 5th Foresight Conference on Molecular Nanotechnology; http://www.foresight.org/Conferences/MNT05/Abstracts/Meyeabst.html.
Saito et al., "A 1M Synapse Self-Learning Digital Neural Network Chip," ISSCC, pp. 6.5-1 to 6.5-10, IEEE 1998.
Espejo. et a;/. "A 16 + 16 Cellular Neural Network Chip for Connected Component Detection," Jun. 30, 1999; http://www.imse.cnm.csic.es/Chipcat/espejo/chip-2.pdf.
Pati et al., "Neural Networks for Tactile Perception," Systems research Center and Dept. of Electrical Engineering, University of Maryland and U.S Naval Research Laboratory. 1987; http://www.isr.umd.edu/TechReports/ISR/1987.TR_87-123.TR_87-123.phtml.
Osamu Fujita, "Statistical estimation of the number of hidden units for feedforward neural networks," Neural Networks 11 (1998), pp. 851-859.
Abraham Harte, "Liquid Crystals Allow Large-Scale Alignment of Carbon Nanotubes," CURJ (Caltech Undergraduate Research Journal), Nov. 2001, vol. 1, No. 2, pp. 44-49.
"Quantum-Dot Arrays for Computation," ORNL Review vol. 34, No. 2, 2001, pp. 1-5 http://www.ornlgov/ORNLReview/v34_2_01/arrays.htm.
Jabri, M.A. et al., "Adaptive Analog VLSI Neural Systems," Chapman & Hall, London SE1 8HN, UK, 1996, pp. 92-95.
Lipson et al., "Automatic Design and Manufacture of Robotic Lifeforms," NATURE, vol. 406, Aug. 31, 2000, pp. 974-978.
Kunitoshi Yamamoto et al., "Rapid Communication Orientation and Purification of Carbon Nanotubes Using AC Electrophoresis", J. Phys. D. Appl. Phys 31 (1998) L34-L36.
E.S. Snow, et al., "Random networks of carbon nanotubes as electronic material", Applied Physics Letters, vol. 82, No. 12, Mar. 31, 2003, pp. 2145-2147.
R. Martel, et al., "Ambipolar Electrical Transport in Semiconducting Single-Wall Carbon Nanotubes," Physical Review Letters, vol. 87, No. 25, Dec. 17, 2001m oo, 256805-1 to 256805-4.
S. Heinze, et al., "Carbon Nanotubes as Schottky Barrier Transistors", vol. 89, No. 10, Sep. 2, 2002, pp. 106801-1 to 106801-4.
M. Dubson, et al., "Measurement of the conductivity exponent in two-dimensional percolating networks: square lattice versus random-void continuum", Physical Review B, vol. 32, No. 11, Dec. 1, 1985, pp. 7621-7623.
D.J. Frank, et al., "Highly efficient algorithm for percolative transport studies in two dimensions", Physical Review B, vol. 37, No. 1, Jan. 1, 1988, pp. 302-307.
Uma R. Karmarkar, et al, "Mechanisms and significance of spike-timing dependent plasticity," Biol. Cybern. 87, 373-382, Jan. 28, 2002.
Uma R. Karmarkar, et al., "A Model of Spike-Timing Dependent Plasticity: One or Two Coincidence Detectors?" J. Neurophysiol, vol. 88, pp. 507-513, Jul. 2002.
M.C.W. van Rossum, et al., "Stable Hebbian Learning from Spkke-Timing-Dependent Plasticity", The Journal of Neuroscience, Dec. 1, 2003, 20(23), pp. 8812-8821.
Xiaohui Xie, et al., "Spike-based learning rules and stabilization of persistent neural activity."
Nace L. Golding, et al., "Dendritic spikes as a mechanism for cooperative long-term potentiation", NATURE, vol. 418, Jul. 18, 2002, pp. 326-330.
Ozgur Turel, et al., "Possible nanoelectronic implementation of neuromorphic networks", Dept. o f Physics and Astronomy, Stony Book University.

V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3; Sep. 9, 2003; American Chemical Society, pp. 1379-1382.
J.M. Tour, et al., "NanoCell Electronic Memories," J.Am. Chem.Soc. 2003, 125, p. 13279-13283.
J. Zaumseil, et al., "Three-Dimensional and Multilayer Nanostructures Formed by Nanotransfer Printing," Nano Letters, 2003, vol. 3, No. 9; Jul. 31, 2003, Amerian Chemical Society, pp. 1223-1227.
Charles D. Schaper, "Patterned Transfer of Metallic Thin Film Nanostructures by Water-Soluble Polymer Templates," Nano Letters, 2003, vol. 3, No. 9; Jul. 26, 2003, American Chemical Society, pp. 1305-1309.
C.A. Dyke, et al., "Unbundled and Highly Functionalized Carbon Nanotubes from Aqueous Reactions," Nano Letters, 2003, vol. 3, No. 9; Aug. 19, 2003, American Chenical Society, pp. 1215-1218.
J. Chung, et al., "Nanoscale Gap Fabrication by Carbon Nanotube-Extracted Lithography (CEL)", Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003, American Chemical Society, pp. 1029-1031.
O. Harnack, et al., "Rectifying Behavior of Electrically Aligned ZnO Nanorods," Nano Letters, 2003, vol. 3, No. 8; Jun. 24, 2003, American Chemical Society, pp. 1097-1101.
M. S. Kumar, et al., "Influence of electric field type on the assembly of single walled carbon nanotubes," Chemical Physics Letters 393 (2004), Dec. 2, 2003; pp. 235-239.
S.W. Lee, et al., "Dielectrophoresis and electrohydrodynamics-mediated fluidic assembly of silicon resistors," Applied Physics Letters, vol. 83, No. 18, Nov. 3, 2003, pp. 3833-3835.
R. Krupke, et al., "Simultaneous Deposition of Metallic Bundles of Single-walled Carbon Nanotubes Using Ac-dielectrophoresis," Nano Letters, 2003, vol. 3, No. 8, Jul. 9, 2003; American Chemical Society, pp. 1019-1023.
K. Bradley, et al., "Flexible Nanotube Electronics," Nano Letters, 2003, vol. 3, No. 10; Aug. 9, 2003, American Chemical Society, pp. 1353-1355.
T.B. Jones, "Frequency-dependent orientation of isolated particle chains," Journal of Electrostatics, 25 (1990), Elsevier Science Publishers, pp. 231-244.
L.A. Hagahara, "Directed placement of suspended carbon nanotubes for nanometer-scale assembly," Applied Physics Letters, vol. 80, No. 20, May 20, 2003; pp. 3826-3828.
A. Bezryadin, et al., "Electrostatic trapping of single conducting nanoparticles between electrodes," Applied Physics Letters, 71 (9), Sep. 1, 1997, pp. 1273-1275.
S. Suzuki, et al., "Quantitative Analysis of DNA Orientation in Stationary AC Electric Fields Using Fluorescence Anistropy," IEEE Transactions of Industry Applications, vol. 34, No. 1, Jan./Feb. 1998, pp. 75-83.
Phaedon Avouris, "Molecular Electronics with Carbon Nanotubes," Accounts of Chemical Research, vol. 35, No. 12; Jul. 31, 2002, pp. 1025-1034.
X. Liu, et al., "Electric-Field Induced Accumulation and Alignment of Carbon Nanotubes," 2002 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp. 31-34.
R. Krupke, et al., "Contacting single bundles of carbon nanotubes with alternating electric fields," Appl. Phys. A. 76, Oct. 28, 2002, pp. 397-400.
M. Law, et al., "Photochemical Sensing of NO_2 with SnO^2 Nanoribbon Nanosensors at Room Temperature," Angew, Chem. 2002, 114m Br, 13m oo, 2511-2514.
J. Tour, et al., "Nanocell Logic Gates for Molecular Computing," IEEE Transactions on Nanotechnology, vol. 1, No. 2, Jun. 2002, pp. 100-109.
A. Leonardi, et al., "Simulation methodology for dielectrophoresis in microelectronic Lab-on-a-chip," Modeling and Simulation of Microsystems 2002, pp. 96-99.
J. Chung, et al., "Nanoscale Gap Fabrication and Integration of Carbon Nanotubes by Micromachining," Solid-State Sensor, Actuator and Microsystems Workshop, Jun. 2-6 2003; Hilton Head Island, South Carolina, pp. 161-164.
L. Zheng, et al., "Towards Single Molecule Manipulation with Dielectrophoresis Using Nanoelectrodes," IEEE-NANO 2003, Aug. 12-14, 2003, Moscone Convention Center, San Francisco, CA; pp. 437-440, http://ieeenano2003.arc.nasa.gov/progran$_{13}$contents.pdf.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3, No. 10; American Chemical Society; Sep. 8, 2003; oo, 1379-1382.
R. Krupe, "Separation of Metallic from Semiconducting Single-Walled Carbon Nanotubes," Science, vol. 301; Jul. 18, 2003; pp. 344-347.
Wolfgang Maass, "On the Relevance of Time in Neural Computation and Learning," In M. Li and A. Maruoka, editors, *Proc. of the 8th International Conference on Algorithmic Learning Theory in Sendai (Japan)*, vol. 1316 of Lecture Notes in Computer Science, pp. 364-388. Springer (Berlin), 1997.
Wolfgang Maass, "Noisy Spiking Neurons with Temporal Coding have more Computational Power than Sigmoidal Neurons," In M. Mozer, M. I. Jordan, T. Petsche, editors, *Advances in Neural Information Processing Systems*, vol. 9, pp. 211-217. MIT Press (Cambridge), 1977. (pp. 1-13, incuding Appendix).
L. Perrinet, et al., "Emergence of filters from natural scenes in a sparse spike coding scheme," Neurocomputing, 2003, pp. 1-14, http://www.laurent.perrinet.free.fr/publi/perrinet03neurocomputing.pdf.
L. Perrinet, et al., "Coherence detection in a spiking neuron via Hebbian learning," Neurocomputing, 2002, vol. 44-46, No. C., pp 817-822, http://www.laurent.perrinet.free.fr/publi/perrinet02.pdf.
A. Jarosz, et al., "An Introductory Note on Gaussian Correlated Random Matrix," Feb. 21, 2003, pp. 1-20 http://www.if.uj.edu.pl/pl/koloSMP/prace/rdnmatrix.pdf.
K. Bradley, et al., "Influence of Mobile Ions on Nanotube Based FET Devices," Nano Letters, 2003, vol. 3, No. 5; American Chemical Society, Apr. 4, 2003; pp. 639-641.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.
Nanoparticles Get Wired, Dimes Institute, Delft University of Technology, 1997.
A. Bezryadin, *Trapping Single Particle with Nanoelectrodes*, Physics News Graphics, Sep. 1997.
Snow, et al., *Nanofabrication with Proximal Probes*, Proceedings of the IEEE, Apr. 1997.
P. O'Connor, G. Gramegna, P. Rehak, F. Corsi, C. Marzocca, *CMOS Preamplifier with High Linearity and Ultra Low Noise for X-Ray Spectroscopy*, IEEE Transactions on Nuclear Science, vol. 44, No. 3, Jun. 1997, pp. 318-325.

* cited by examiner

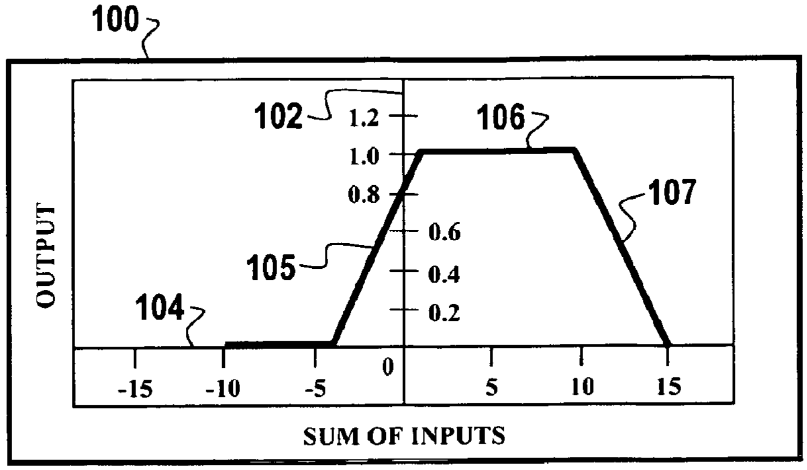


Figure 1

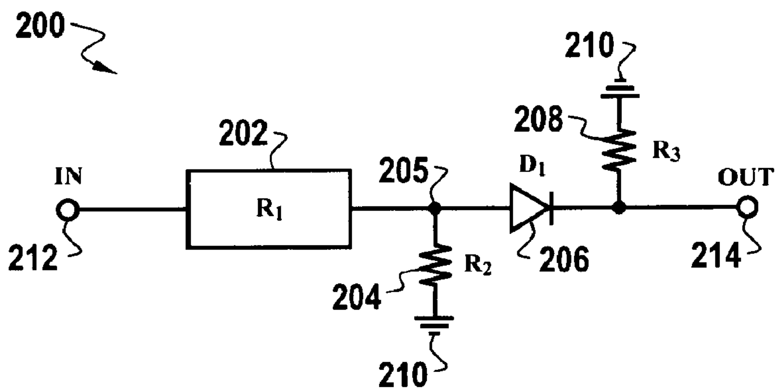


Figure 2

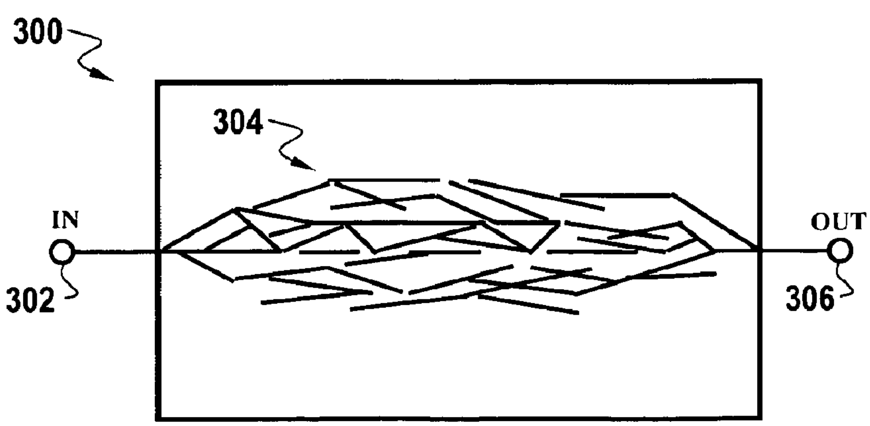


Figure 3

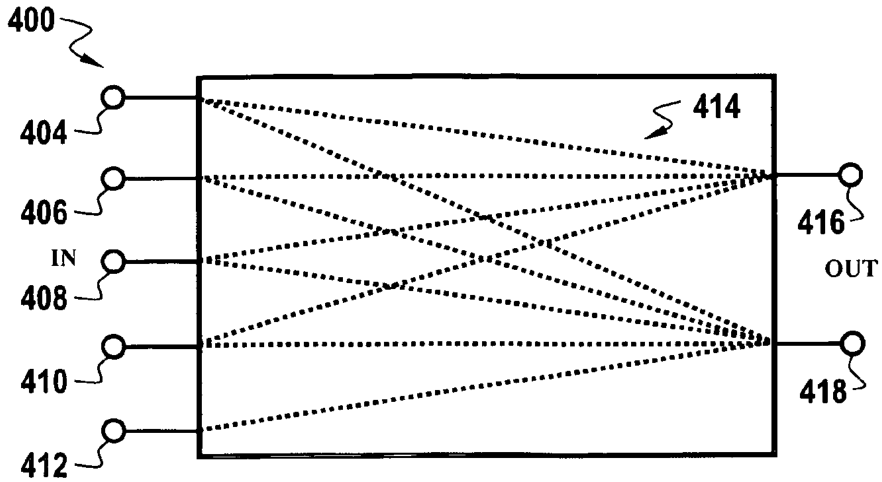


Figure 4

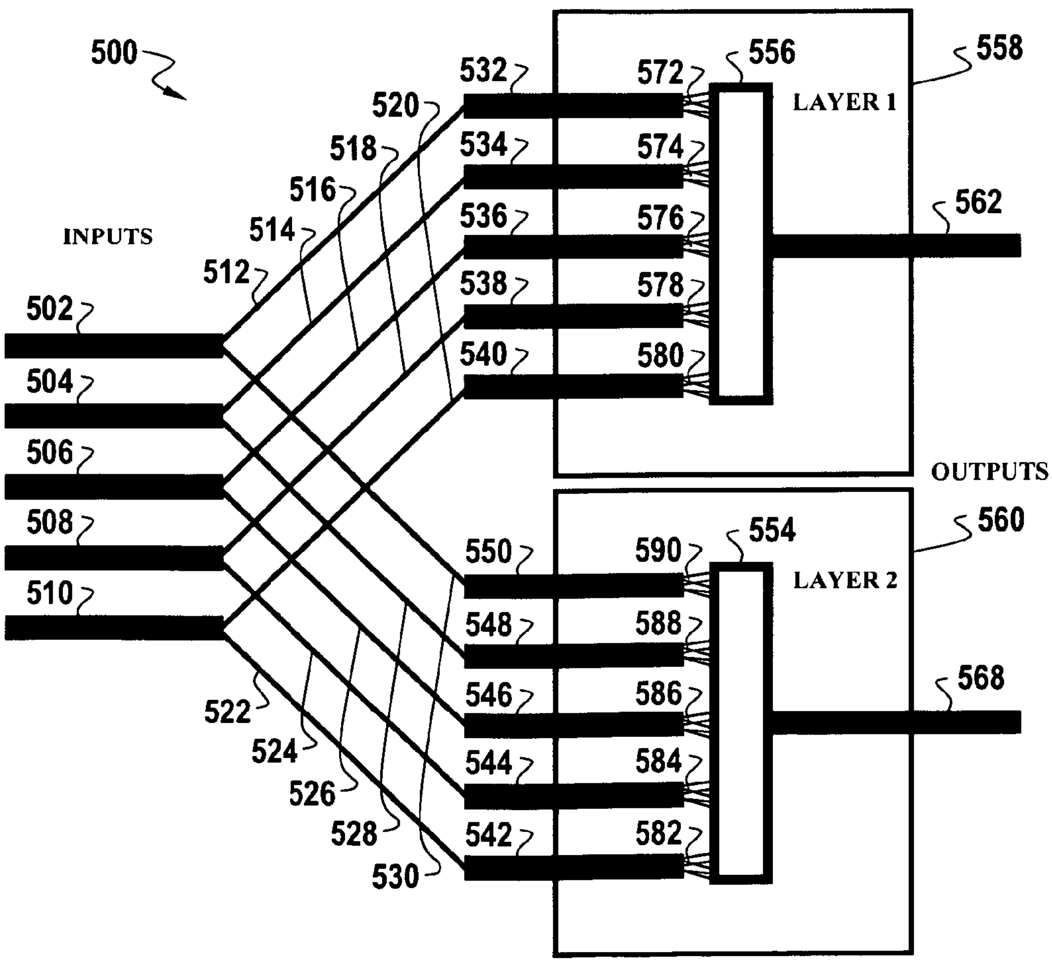


Figure 5

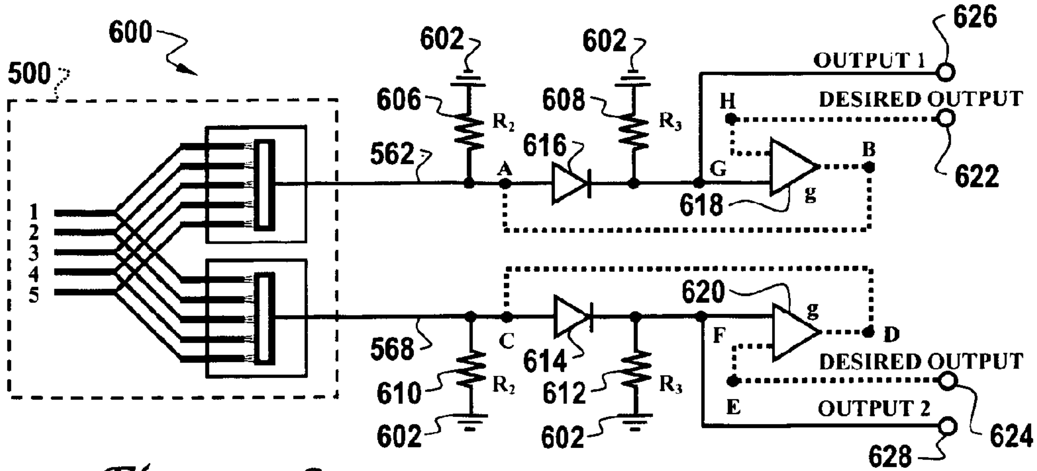
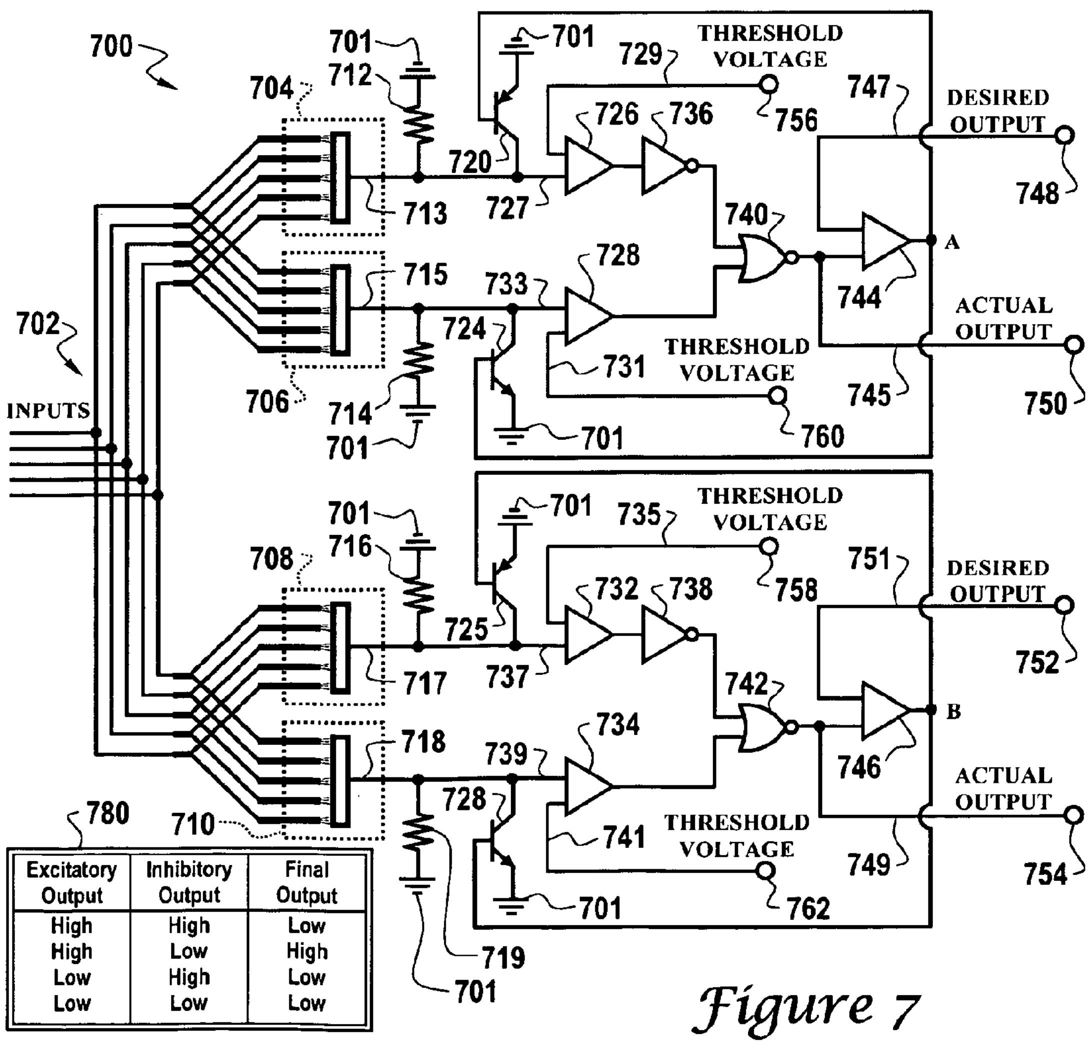


Figure 6



Excitatory Output	Inhibitory Output	Final Output
High	High	Low
High	Low	High
Low	High	Low
Low	Low	Low

Figure 7

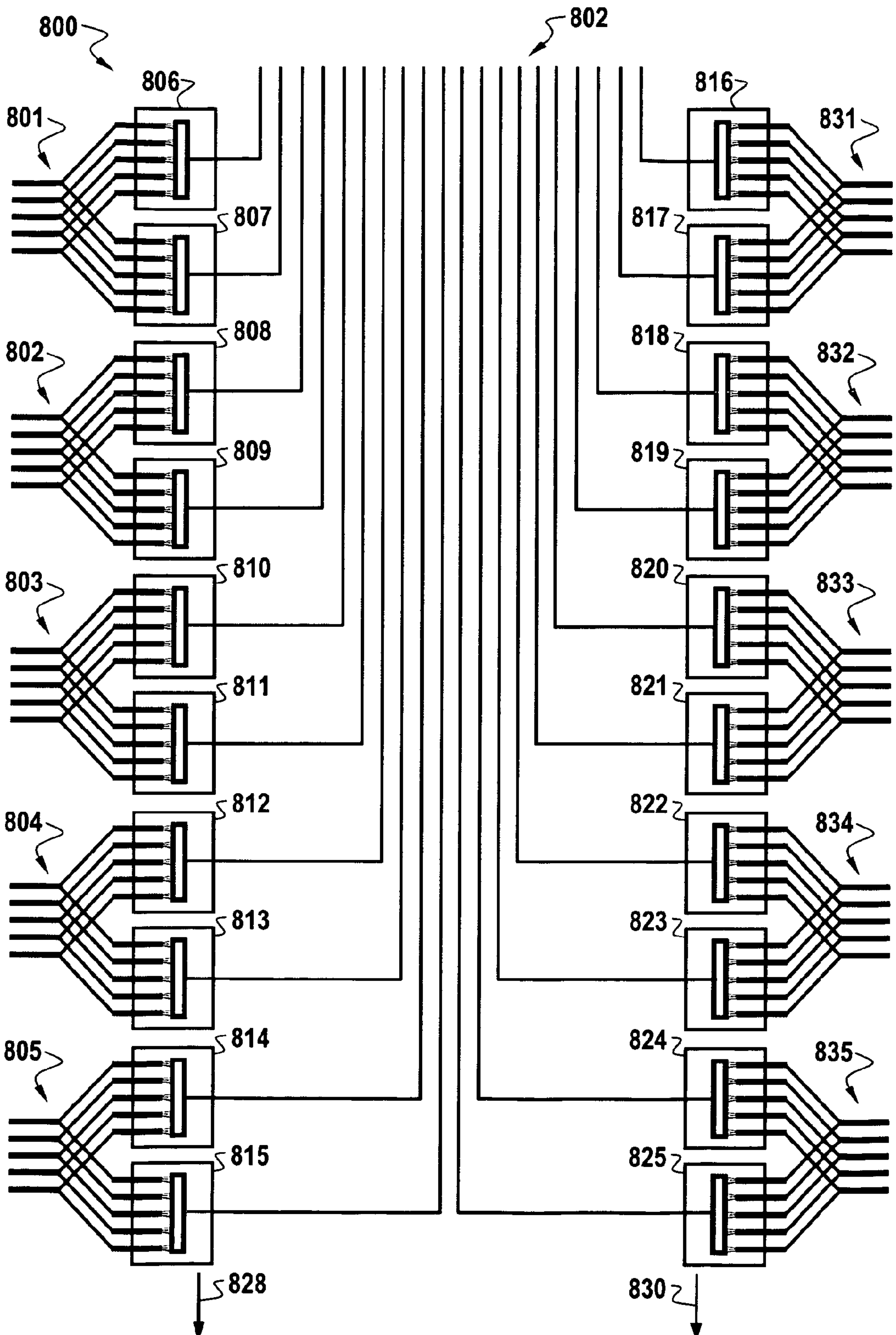

800
802
801
806
807
808
809
810
811
812
813
814
815
802
803
804
805
816
817
818
819
820
821
822
823
824
825
831
832
833
834
835
828
830

Figure 8

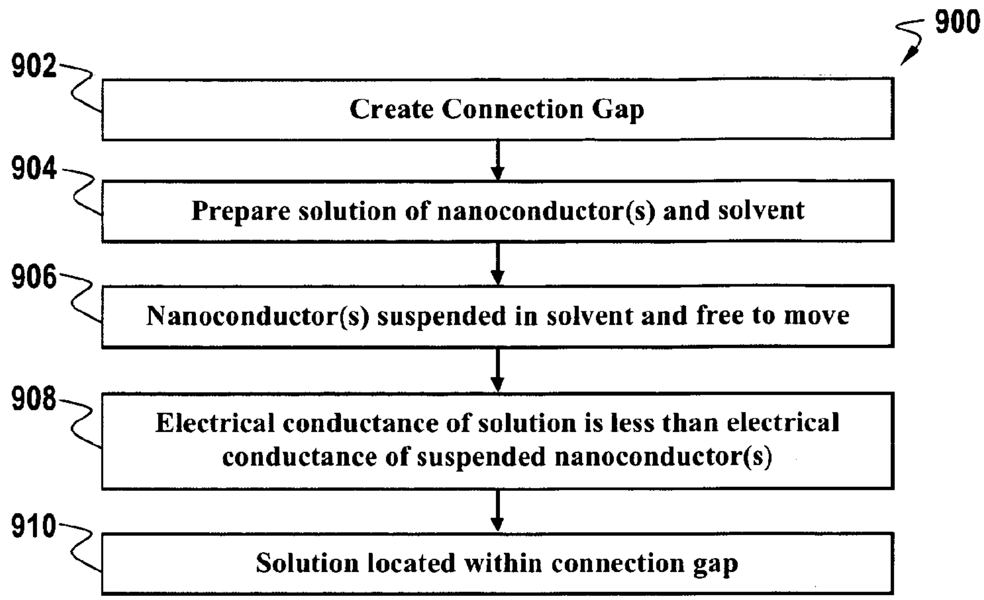


Figure 9

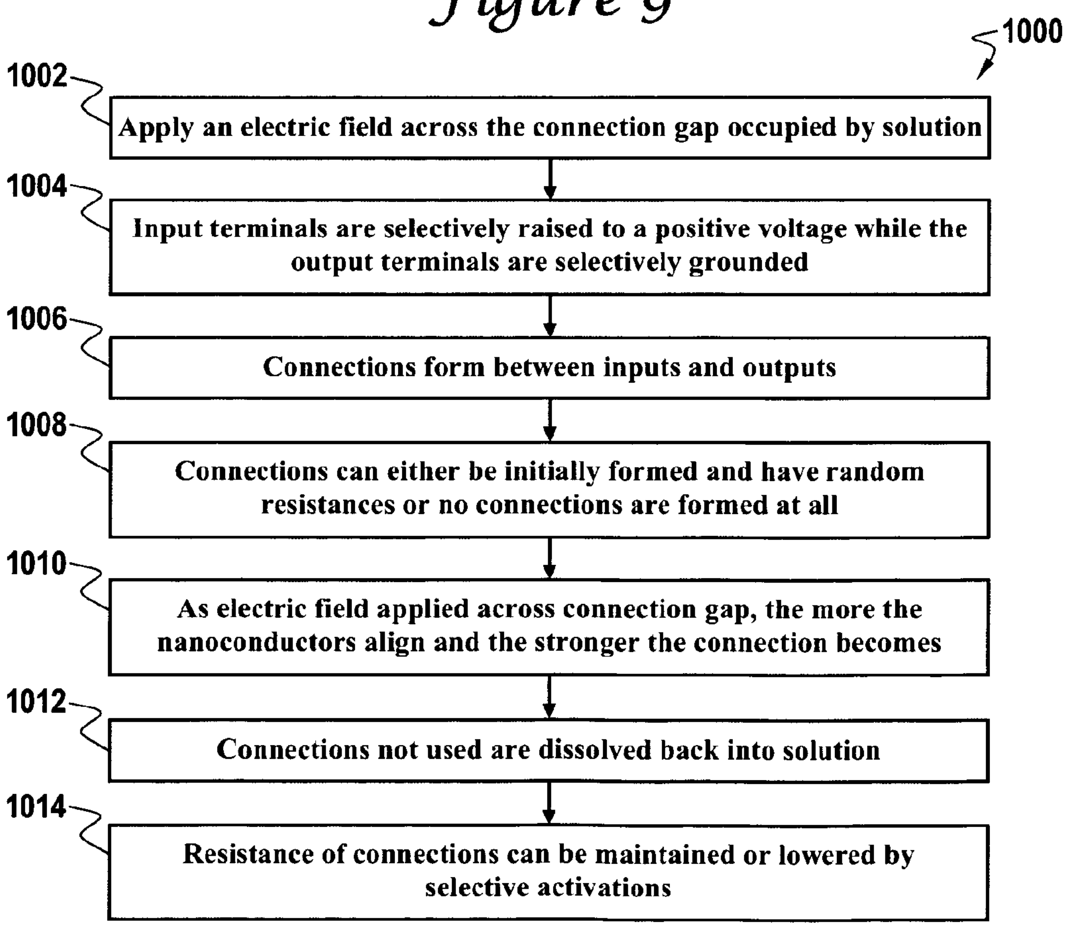


Figure 10

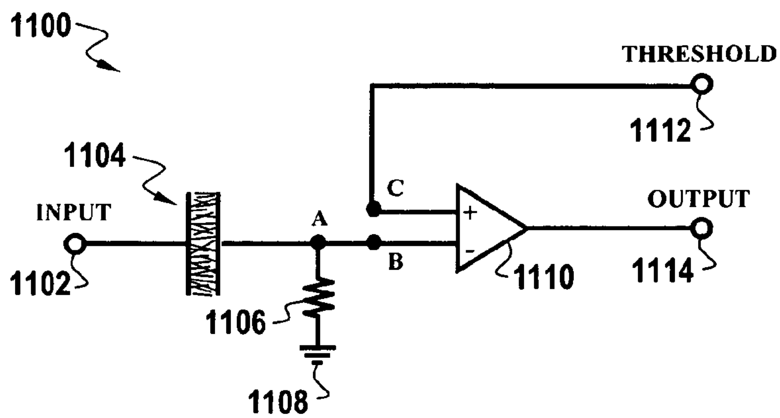


Figure 11

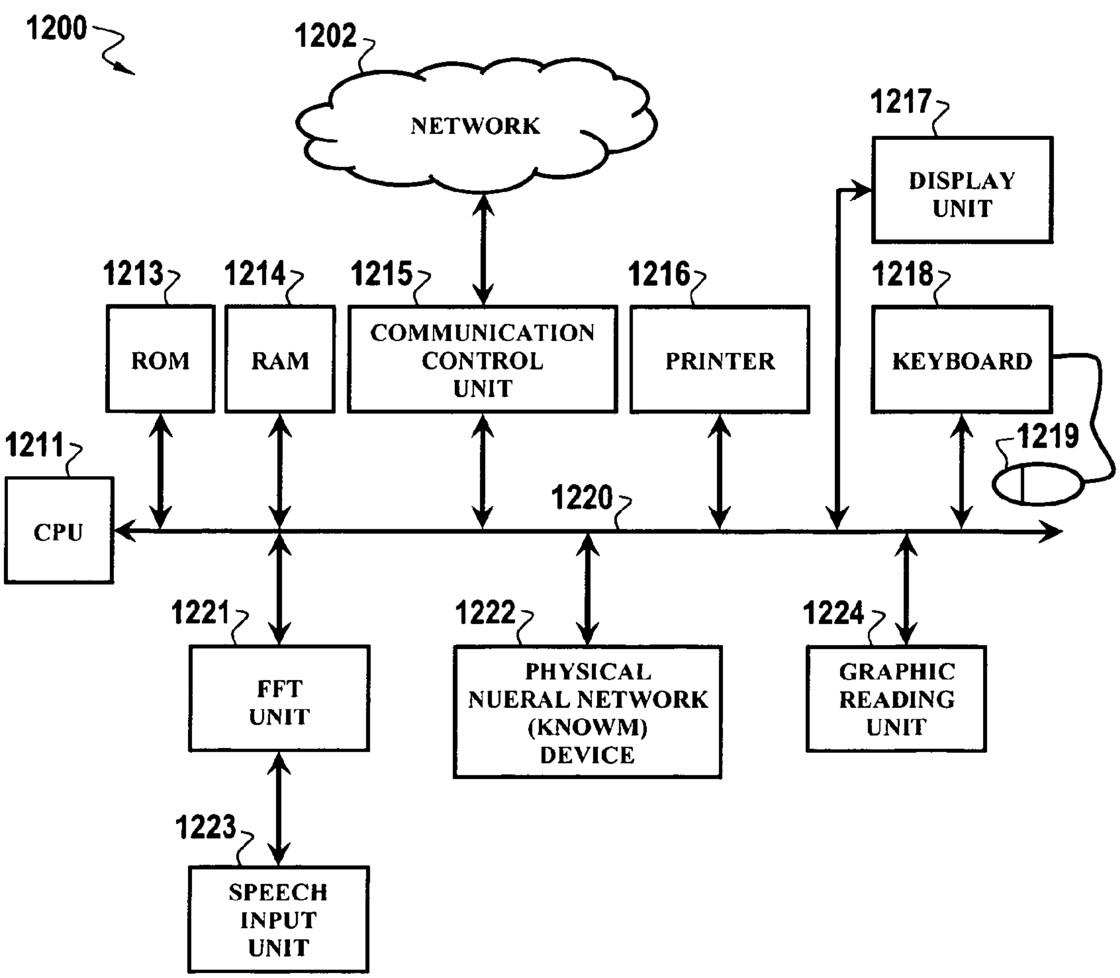


Figure 12

UTILIZED NANOTECHNOLOGY APPARATUS USING A NEUTRAL NETWORK, A SOLUTION AND A CONNECTION GAP

CROSS REFERENCE TO RELATED PATENT APPLICATION

This patent application is a continuation of U.S. patent application Ser. No. 10/095,273 entitled "Physical Neural Network Design Incorporating Nanotechnology," which was filed on Mar. 12, 2002, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

Embodiments generally relate to nanotechnology. Embodiments also relate to nanotechnology-based components. Embodiments additionally relate to artificial neural networks and solutions and solvents for maintaining such networks and components thereof.

BACKGROUND

Neural networks are computational systems that permit computers to essentially function in a manner analogous to that of the human brain. Neural networks do not utilize the traditional digital model of manipulating 0's and 1's. Instead, neural networks create connections between processing elements, which are equivalent to neurons of a human brain. Neural networks are thus based on various electronic circuits that are modeled on human nerve cells (i.e., neurons). Generally, a neural network is an information-processing network, which is inspired by the manner in which a human brain performs a particular task or function of interest. Computational or artificial neural networks are thus inspired by biological neural systems. The elementary building block of biological neural systems is of course the neuron, the modifiable connections between the neurons, and the topology of the network.

In general, artificial neural networks are systems composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. The computational elements, or nodes, are connected via variable weights that are typically adapted during use to improve performance. Thus, in solving a problem, neural net models can explore many competing hypothesis simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights.

In a neural network, "neuron-like" nodes can output a signal based on the sum of their inputs, the output being the result of an activation function. In a neural network, there exists a plurality of connections, which are electrically coupled among a plurality of neurons. The connections serve as communication bridges among of a plurality of neurons coupled thereto. A network of such neuron-like nodes has the ability to process information in a variety of useful ways. By adjusting the connection values between neurons in a network, one can match certain inputs with desired outputs.

Neural networks that have been developed to date are largely software-based. A true neural network (e.g., the human brain) is massively parallel (and therefore very fast computationally) and very adaptable. For example, half of a human brain can suffer a lesion early in its development and not seriously affect its performance. Software simulations are slow because during the learning phase a standard

computer must serially calculate connection strengths. When the networks get larger (and therefore more powerful and useful), the computational time becomes enormous. For example, networks with 10,000 connections can easily overwhelm a computer. In comparison, the human brain has about 100 billion neurons, each of which can be connected to about 5,000 other neurons. On the other hand, if a network is trained to perform a specific task, perhaps taking many days or months to train, the final useful result can be etched onto a piece of silicon and also mass-produced.

A number of software simulations of neural networks have been developed. Because software simulations are performed on conventional sequential computers, however, they do not take advantage of the inherent parallelism of neural network architectures. Consequently, they are relatively slow.

The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions. Due to the difficulties in building such highly interconnected processors, the currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and built to implement one particular neural network architecture and are not easily, if at all, reconfigurable to implement different architectures. A true physical neural network (i.e., artificial neural network) chip, for example, has not yet been designed and successfully implemented.

The problem with pure hardware implementation of a neural network with technology as it exists today, is the inability to physically form a great number of connections and neurons. On-chip learning can exist, but the size of the network would be limited by digital processing methods and associated electronic circuitry. One of the difficulties in creating true physical neural networks lies in the highly complex manner in which a physical neural network must be designed and built. It is believed that solutions to creating a true physical and artificial neural network lie in the use of nanotechnology and the implementation of analog variable connections.

The term "Nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

Present chip technology is also limiting when wires need to be crossed on a chip. For the most part, the design of a computer chip is limited to two dimensions. Each time a circuit must cross another circuit, another layer must be added to the chip. This increases the cost and decreases the speed of the resulting chip. A number of alternatives to standard silicon based complementary metal oxide semiconductor ("CMOS") devices have been proposed. The common goal is to produce logic devices on a nanometer scale.

Such dimensions are more commonly associated with molecules than integrated circuits.

Integrated circuits and electrical components thereof, which can be produced at a molecular and nanometer scale, include devices such as carbon nanotubes and nanowires, which essentially are nanoscale conductors ("nanoconductors"). Nanocondcutors are tiny conductive tubes (i.e., hollow) or wires (i.e., solid) with a very small size scale (e.g., 1.0–100 nanometers in diameter and hundreds of microns in length). Their structure and fabrication have been widely reported and are well known in the art. Carbon nanotubes, for example, exhibit a unique atomic arrangement, and possess useful physical properties such as one-dimensional electrical behavior, quantum conductance, and ballistic electron transport.

Carbon nanotubes are among the smallest dimensioned nanotube materials with a generally high aspect ratio and small diameter. High-quality single-walled carbon nanotubes can be grown as randomly oriented, needle-like or spaghetti-like tangled tubules. They can be grown by a number of fabrication methods, including chemical vapor deposition (CVD), laser ablation or electric arc growth. Carbon nanotubes can be grown on a substrate by catalytic decomposition of hydrocarbon containing precursors such as ethylene, methane, or benzene. Nucleation layers, such as thin coatings of Ni, Co, or Fe are often intentionally added onto the substrate surface in order to nucleate a multiplicity of isolated nanotubes. Carbon nanotubes can also be nucleated and grown on a substrate without a metal nucleating layer by using a precursor including one or more of these metal atoms. Semiconductor nanowires can be grown on substrates by similar processes.

Based on the foregoing, it is believed that a physical neural network which incorporates nanotechnology is a solution to the problems encountered by prior art neural network solutions. It is believed that a true physical neural network can be designed and constructed without relying on computer simulations for training, or relying on standard digital (binary) memory to store connections strengths. Additionally, it is believed that a variable resistor apparatus can be constructed based on nanotechnology and utilized either as an individual component for variable resistance purposes, or in association with physical neural networks, including artificial neurons and components thereof as described herein.

BRIEF SUMMARY

The following summary is provided to facilitate an understanding of some of the innovative features unique to the embodiments, and is not intended to be a full description. A full appreciation of the various aspects of the embodiments can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

It is, therefore, one aspect of the present invention to provide for a solution for maintaining nanoparticles utilized in a physical neural network.

It is another aspect of the present invention to provide a physical neural network, which can be formed from a plurality of interconnected nanoconnections or nanoconnectors.

It is a further aspect of the present invention to provide neuron like nodes, which can be formed and implemented utilizing nanotechnology;

It is also an aspect of the present invention to provide a physical neural network that can be formed from one or more neuron-like nodes.

It is yet a further aspect of the present invention to provide a physical neural network, which can be formed from a plurality of nanoconductors, such as, for example, nanowires and/or nanotubes.

The above and other aspects can be achieved as is now described. A physical neural network based on nanotechnology is disclosed herein, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, connected to a plurality of interconnected nanoconnections. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes can be strengthened or weakened according to an application of an electric field. Alignment has also been observed with a magnetic field, but electric fields are generally more practical. Note that the connection network is generally associated with one or more of the neuron-like nodes.

The output signal is generally based on a threshold below which the output signal is not generated and above which the output signal is generated. The transition from zero output to high output need not necessarily be abrupt or non linear. The connection network comprises a number of layers of nanoconnections, wherein the number of layers is generally equal to a number of desired outputs from the connection network. The nanoconnections are formed without influence from disturbances resulting from other nanoconnections thereof. Such nanoconnections may be formed from an electrically conducting material. The electrically conducting material can be selected such that a dipole is induced in the electrically conducting material in the presence of an electric field. Such a nanoconnection may comprise a nanoconductor.

The connection network itself may comprise a connection network structure having a connection gap formed therein, and a solution located within the connection gap, such that the solution comprises a solvent or suspension and one or more nanoconductors. Preferably, a plurality of nanoconductors is present in the solution (i.e., mixture). Note that such a solution may comprise a liquid and/or gas. An electric field can then be applied across the connection gap to permit the alignment of one or more of the nanoconductors within the connection gap. The nanoconductors can be suspended in the solvent, or can lie at the bottom of the connection gap on the surface of the chip. Studies have shown that nanotubes can align both in the suspension and/or on the surface of the gap. The electrical conductance of the mixture is less than the electrical conductance of the nanoconductors within the solution.

The nanoconductors within the connection gap thus experience an increased alignment in accordance with an increase in the electric field applied across the connection gap. Thus, nanoconnections of the neuron-like node that are utilized most frequently by the neuron-like node become stronger with each use thereof. The nanoconnections that are utilized least frequently become increasingly weak and eventually dissolve back into the solution. The nanoconnections may or may not comprise a resistance, which can be raised or lowered by a selective activation of a nanoconnection. They can be configured as nanoconductors such as, for example, a nanotube or nanowire. An example of a nanotube, which may be implemented in accordance with the invention described herein, is a carbon nanotube, nanowire and/or other nanoparticle. Additionally, such nanocon-

nections may be configured as a negative connection associated with the neuron-like node.

In general, an apparatus for maintaining components in physical neural network formed utilizing nanotechnology is described herein. A connection gap can be formed between two terminals. A solution comprising a melting point at approximately room temperature can be provided, wherein the solution is maintained in the connection gap and comprises a plurality of nanoparticles forming nanoconnections thereof having connection strengths thereof, wherein the solution and the connection gap are adapted for use with a physical neural network formed utilizing nanotechnology, such when power is removed from the physical neural network, the solution freezes, thereby locking into place the connection strengths.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a graph illustrating a typical activation function that can be implemented in accordance with one embodiment;

FIG. **2** illustrates a schematic diagram illustrating a diode configuration as a neuron, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanowires between two electrodes, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections formed between two electrodes, in accordance with a preferred embodiment;

FIG. **4** illustrates a block diagram illustrating a plurality of connections between inputs and outputs of a physical neural network, in accordance with a preferred embodiment;

FIG. **5** illustrates a schematic diagram of a physical neural network that can be created without disturbances, in accordance with a preferred embodiment;

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **8** illustrates a schematic diagram of a chip layout for a connection network that may be implemented in accordance with an alternative embodiment;

FIG. **9** illustrates a flow chart of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment;

FIG. **10** illustrates a flow chart of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment;

FIG. **11** illustrates a schematic diagram of a circuit illustrating temporal summation within a neuron, in accordance with a preferred embodiment; and

FIG. **12** illustrates a block diagram illustrating a pattern recognition system, which may be implemented with a physical neural network device, in accordance with a preferred embodiment.

DETAILED DESCRIPTION

The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate one or more embodiments.

The physical neural network described and disclosed herein is different from prior art forms of neural networks in that the disclosed physical neural network does not require a computer simulation for training, nor is its architecture based on any current neural-network hardware device. The design of the physical neural network described herein with respect to particular embodiments is actually quite "organic". Such a physical neural network is generally fast and adaptable, no matter how large such a physical neural network becomes. The physical neural network described herein can be referred to generically as a Knowm. The terms "physical neural network" and "Knowm" can be utilized interchangeably to refer to the same device, network, or structure.

Network orders of magnitude larger than current VSLI neural networks can be built and trained with a standard computer. One consideration for a Knowm is that it must be large enough for its inherent parallelism to shine through. Because the connection strengths of such a physical neural network are dependant on the physical movement of nanoconnections thereof, the rate at which a small network can learn is generally very small and a comparable network simulation on a standard computer can be very fast. On the other hand, as the size of the network increases, the time to train the device does not change. Thus, even if the network takes a full second to change a connection value a small amount, if it does the same to a billion connections simultaneously, then its parallel nature begins to express itself.

A physical neural network (i.e., a Knowm) must have two components to function properly. First, the physical neural network must have one or more neuron-like nodes that sum a signal and output a signal based on the amount of input signal received. Such a neuron-like node is generally nonlinear in its output. In other words, there should be a certain threshold for input signals, below which nothing is output and above which a constant or nearly constant output is generated or allowed to pass. This is a very basic requirement of standard software-based neural networks, and can be accomplished by an activation function. The second requirement of a physical neural network is the inclusion of a connection network composed of a plurality of interconnected connections (i.e., nanoconnections). Such a connection network is described in greater detail herein.

FIG. **1** illustrates a graph **100** illustrating a typical activation function that can be implemented in accordance with one embodiment. Note that the activation function need not be non-linear, although non-linearity is generally desired for learning complicated input-output relationships. The activation function depicted in FIG. **1** comprises a linear function, and is shown as such for general edification and illustrative purposes only. As explained previously, an activation function may also be non-linear.

As illustrated in FIG. **1**, graph **100** includes a horizontal axis **104** representing a sum of inputs, and a vertical axis **102** representing output values. A graphical line **106** indicates threshold values along a range of inputs from approximately −10 to +10 and a range of output values from approximately 0 to 1. As more neural networks (i.e., active inputs) are established, the overall output as indicated at line **105** climbs until the saturation level indicated by line **106** is attained. If a connection is not utilized, then the level of output (i.e., connection strength) begins to fade until it is revived. This phenomenon is analogous to short term memory loss of a human brain. Note that graph **100** is presented for generally illustrative and edification purposes only and is not considered a limiting feature of the embodiments.

In a Knowm network, the neuron-like node can be configured as a standard diode-based circuit, the diode being the most basic semiconductor electrical component, and the signal it sums may be a voltage. An example of such an arrangement of circuitry is illustrated in FIG. **2**, which generally illustrates a schematic diagram illustrating a diode-based configuration as a neuron **200**, in accordance with a preferred embodiment. Those skilled in the art can appreciate that the use of such a diode-based configuration is not considered a limitation of the embodiments, but merely represents one potential arrangement in which the embodiments may be implemented.

Although a diode may not necessarily be utilized, its current versus voltage characteristics are non-linear when used with associated resistors and similar to the relationship depicted in FIG. **1**. The use of a diode as a neuron is thus not a limiting feature, but is only referenced herein with respect to a preferred embodiment. The use of a diode and associated resistors with respect to a preferred embodiment simply represents one potential "neuron" implementation. Such a configuration can be said to comprise an artificial neuron. It is anticipated that other devices and components may be utilized instead of a diode to construct a physical neural network and a neuron-like node (i.e., artificial neuron), as indicated here.

Thus, neuron **200** comprises a neuron-like node that may include a diode **206**, which is labeled D_1, and a resistor **204**, which is labeled R_2. Resistor **204** is connected to a ground **210** and an input **205** of diode **206**. Additionally, a resistor **202**, which is represented as a block and labeled R_1 can be connected to input **205** of diode **206**. Block **202** includes an input **212**, which comprises an input to neuron **200**. A resistor **208**, which is labeled R_3, is also connected to an output **214** of diode **206**. Additionally, resistor **208** is coupled to ground **210**. Diode **206** in a physical neural network is analogous to a neuron of a human brain, while an associated connection formed thereof, as explained in greater detail herein, is analogous to a synapse of a human brain.

As depicted in FIG. **2**, the output **214** is determined by the connection strength of R_1 (i.e., resistor **202**). If the strength of R_1's connection increases (i.e., the resistance decreases), then the output voltage at output **214** also increases. Because diode **206** conducts essentially no current until its threshold voltage (e.g., approximately 0.6V for silicon) is attained, the output voltage will remain at zero until R_1 conducts enough current to raise the pre-diode voltage to approximately 0.6V. After 0.6V has been achieved, the output voltage at output **214** will increase linearly. Simply adding extra diodes in series or utilizing different diode types may increase the threshold voltage.

An amplifier may also be added to the output **214** of diode **206** so that the output voltage immediately saturates at the diode threshold voltage, thus resembling a step function, until a threshold value and a constant value above the threshold is attained. R_3 (i.e., resistor **208**) functions generally as a bias for diode **206** (i.e., D_1) and should generally be about 10 times larger than resistor **204** (i.e., R_2). In the circuit configuration illustrated in FIG. **2**, R_1 can actually be configured as a network of connections composed of many inter-connected conducting nanowires (i.e., see FIG. **3**). As explained previously, such connections are analogous to the synapses of a human brain.

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections **304** formed between two electrodes, in accordance with a preferred embodiment. Nanoconnections **304** (e.g., nanoconductors) depicted in FIG. **3** are generally located between input **302** and output **306**. The network of nanoconnections depicted in FIG. **3** can be implemented as a network of nanoconductors. Examples of nanoconductors include devices such as, for example, nanowires, nanotubes, and nanoparticles.

Nanoconnections **304**, which are analogous to the synapses of a human brain, are preferably composed of electrical conducting material (i.e., nanoconductors). It should be appreciated by those skilled in the art that such nanoconductors can be provided in a variety of shapes and sizes without departing from the teachings herein. For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles.

For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles. For example, each nanoconnection within a connection network may be formed from as a chain of carbon nanoparticles. In "Self-assembled chains of graphitized carbon nanoparticles" by Bezryadin et al., Applied Physics Letters, Vol. 74, No. 18, pp. 2699–2701, May 3, 1999, for example, a technique is reported, which permits the self-assembly of conducting nanoparticles into long continuous chains. Thus, nanoconductors which are utilized to form a physical neural network (i.e., Knowm) could be formed from such nanoparticles. It can be appreciated that the Bezryadin et al is referred to herein for general edification and illustrative purposes only and is not considered to limit the embodiments.

It can be appreciated that a connection network as disclosed herein may be composed from a variety of different types of nanoconductors. For example, a connection network may be formed from a plurality of nanoconductors, including nanowires, nanotubes and/or nanoparticles. Note that such nanowires, nanotubes and/or nanoparticles, along with other types of nanoconductors can be formed from materials such as carbon or silicon. For example, carbon nanotubes may comprise a type of nanotube that can be utilized in accordance with one or more embodiments.

As illustrated in FIG. **3**, nanoconnections **304** comprise a plurality of interconnected nanoconnections, which from this point forward, can be referred to generally as a "connection network." An individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire, a nanotube, nanoparticles(s), or any other nanoconducting structures. Nanoconnections **304** may comprise a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections **304** may be formed from a plurality of interconnected nanoparticles. A connection network is thus not one connection between two electrodes, but a plurality of connections between inputs and outputs. Nanotubes, nanowires, nanoparticles and/or other nanoconducting structures may be utilized, of course, to construct nanoconnections **304** between input **302** and input **306**. Although a single input **302** and a single input **306** is depicted in FIG. **3**, it can be appreciated that a plurality of inputs and a plurality of outputs may be implemented in accordance with the embodiments, rather than simply a single input **302** or a single output **306**.

FIG. **4** illustrates a block diagram illustrating a plurality of nanoconnections **414** between inputs **404**, **406**, **408**, **410**, **412** and outputs **416** and **418** of a physical neural network, in accordance with a preferred embodiment. Inputs **404**, **406**, **408**, **410**, and **412** can provide input signals to connections **414**. Output signals can then be generated from

connections **414** via outputs **416** and **418**. A connection network can therefore be configured from the plurality of connections **414**. Such a connection network is generally associated with one or more neuron-like nodes.

The connection network also comprises a plurality of interconnected nanoconnections, wherein each nanoconnection thereof is strengthened or weakened according to an application of an electric field. A connection network is not possible if built in one layer because the presence of one connection can alter the electric field so that other connections between adjacent electrodes could not be formed. Instead, such a connection network can be built in layers, so that each connection thereof can be formed without being influenced by field disturbances resulting from other connections. This can be seen in FIG. **5**.

FIG. **5** illustrates a schematic diagram of a physical neural network **500** that can be created without disturbances, in accordance with a preferred embodiment. Physical neural network **500** is composed of a first layer **558** and a second layer **560**. A plurality of inputs **502**, **504**, **506**, **508**, and **510** are respectively provided to layers **558** and **560** respectively via a plurality of input lines **512**, **514**, **516**, **518**, and **520** and a plurality of input lines **522**, **524**, **526**, **528**, and **530**. Input lines **512**, **514**, **516**, **518**, and **520** are further coupled to input lines **532**, **534**, **536**, **538**, and **540** such that each line **532**, **534**, **536**, **538**, and **540** is respectively coupled to nanoconnections **572**, **574**, **576**, **578**, and **580**. Thus, input line **532** is connected to nanconnections **572**. Input line **534** is connected to nanoconnections **574**, and input line **536** is connected to nanoconnections **576**. Similarly, input line **538** is connected to nanconnections **578**, and input line **540** is connected to nanoconnections **580**.

Nanconnections **572**, **574**, **576**, **578**, and **580** may comprise nanoconductors such as, for example, nanotubes and/or nanowires. Nanoconnections **572**, **574**, **576**, **578**, and **580** thus comprise one or more nanoconductors. Additionally, input lines **522**, **524**, **526**, **528**, and **530** are respectively coupled to a plurality of input lines **542**, **544**, **546**, **548** and **550**, which are in turn each respectively coupled to nanoconnections **582**, **584**, **586**, **588**, and **590**. Thus, for example, input line **542** is connected to nanoconnections **582**, while input line **544** is connected to nanoconnections **584**. Similarly, input line **546** is connected to nanoconnections **586** and input line **548** is connected to nanoconnections **588**. Additionally, input line **550** is connected to nanconnections **590**. Box **556** and **554** generally represent simply the output and are thus illustrated connected to outputs **562** and **568**. In other words, outputs **556** and **554** respectively comprise outputs **562** and **568**. The aforementioned input lines and associated components thereof actually comprise physical electronic components, including conducting input and output lines and physical nanoconnections, such as nanotubes and/or nanowires.

Thus, the number of layers **558** and **560** equals the number of desired outputs **562** and **568** from physical neural network **500**. In the previous two figures, every input was potentially connected to every output, but many other configurations are possible. The connection network can be made of any electrically conducting material, although the physics of it requires that they be very small so that they will align with a practical voltage. Carbon nanotubes or any conductive nanowire can be implemented in accordance with the physical neural network described herein. Such components can form connections between electrodes by the presence of an electric field. For example, the orientation and purification of carbon nanotubes has been demonstrated using ac electrophoresis in isopropyl alcohol, as indicated in "Orientation and purification of carbon nanotubes using ac electrophoresis" by Yamamoto et al., J. Phys. D: Applied Physics, 31 (1998), 34–36. Additionally, an electric-field assisted assembly technique used to position individual nanowires suspended in an electric medium between two electrodes defined lithographically on an SiO_2 substrate is indicated in "Electric-field assisted assembly and alignment of metallic nanowires," by Smith et al., Applied Physics Letters, Vol. 77, Num. 9, Aug. 28, 2000. Such references are referred to herein for edification and illustrative purposes only.

The only general requirements for the conducting material utilized to configure the nanoconductors are that such conducting material should preferably conduct electricity, and a dipole should preferably be induced in the material when in the presence of an electric field. Alternatively, the nanoconductors utilized in association with the physical neural network described herein can be configured to include a permanent dipole that is produced by a chemical means, rather than a dipole that is induced by an electric field.

Therefore, it should be appreciated by those skilled in the art that a connection network could also be comprised of other conductive particles that may be developed or found useful in the nanotechnology arts. For example, carbon particles (or "dust") may also be used as nanoconductors in place of nanowires or nanotubes. Such particles may include bearings or granule-like particles.

A connection network can be constructed as follows: A voltage is applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture could be made of many things. The only requirements are that the conducting wires must be suspended in the solvent, either dissolved or in some sort of suspension, free to move around; the electrical conductance of the substance must be less than the electrical conductance of the suspended conducting wire; and the viscosity of the substance should not be too much so that the conducting wire cannot move when an electric field is applied.

The goal for such a connection network is to develop a network of connections of just the right values so as to satisfy the particular signal—processing requirement—exactly what a neural network does. Such a connection network can be constructed by applying a voltage across a space occupied by the mixture mentioned. To create the connection network, the input terminals are selectively raised to a positive voltage while the output terminals are selectively grounded. Thus, connections can gradually form between the inputs and outputs. The important requirement that makes the physical neural network functional as a neural network is that the longer this electric field is applied across a connection gap, or the greater the frequency or amplitude, the more nanotubes and/or nanowires and/or particles align and the stronger the connection thereof becomes. Thus, the connections that are utilized most frequently by the physical neural network become the strongest.

The connections can either be initially formed and have random resistances or no connections may be formed at all. By initially forming random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up from scratch. Depending on the rate of connection decay, having initial random connections could prove faster, although not necessarily. The connection network can adapt itself to the requirements of a given situation regardless of the initial state of the connections. Either initial condition will work, as connections that are not used will "dissolve" back into

solution. The resistance of the connection can be maintained or lowered by selective activations of the connection. In other words, if the connection is not used, it will fade away, analogous to the connections between neurons in a human brain. The temperature of the solution can also be maintained at a particular value so that the rate that connections fade away can be controlled. Additionally an electric field can be applied perpendicular to the connections to weaken them, or even erase them out altogether (i.e., as in clear, zero, or reformatting of a "disk").

The nanoconnections may or may not be arranged in an orderly array pattern. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, due to the non-linear nature of neural networks, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network **600** that can be implemented in accordance with an alternative embodiment. Note that in FIGS. **5** and **6**, like parts are indicated by like reference numerals. Thus, physical neural network **600** can be configured, based on physical neural network **500** illustrated in FIG. **5**. In FIG. **6**, inputs **1**, **2**, **3**, **4**, and **5** are indicated, which are respectively analogous to inputs **502**, **504**, **506**, **508**, and **510** illustrated in FIG. **5**. Outputs **562** and **568** are provided to a plurality of electrical components to create a first output **626** (i.e., Output **1**) and a second output **628** (i.e., Output **2**). Output **562** is tied to a resistor **606**, which is labeled R2 and a diode **616** at node A. Output **568** is tied to a resistor **610**, which is also labeled R2 and a diode **614** at node C. Resistors **606** and **610** are each tied to a ground **602**.

Diode **616** is further coupled to a resistor **608**, which is labeled R3, and first output **626**. Additionally, resistor **608** is coupled to ground **602** and an input to an amplifier **618**. An output from amplifier **618**, as indicated at node B and dashed lines thereof, can be tied back to node A. A desired output **622** from amplifier **618** is coupled to amplifier **618** at node H. Diode **614** is coupled to a resistor **612** at node F. Note that resistor **612** is labeled R3. Node F is in turn coupled to an input of amplifier **620** and to second output **628** (i.e., Output **2**). Diode **614** is also connected to second output **628** and an input to amplifier **620** at second output **628**. Note that second output **628** is connected to the input to amplifier **620** at node F. An output from amplifier **620** is further coupled to node D, which in turn is connected to node C. A desired output **624**, which is indicated by a dashed line in FIG. **6**, is also coupled to an input of amplifier **620** at node E.

In FIG. **6**, the training of physical neural network **600** can be accomplished utilizing, for example, op-amp devices (e.g., amplifiers **618** and **620**). By comparing an output (e.g., first output **626**) of physical neural network **600** with a desired output (e.g., desired output **622**), the amplifier (e.g., amplifier **618**) can provide feedback and selectively strengthen connections thereof. For instance, suppose it is desired to output a voltage of +V at first output **626** (i.e., Output **1**) when inputs **1** and **4** are high. When inputs **1** and **4** are taken high, also assume that first output **626** is zero. Amplifier **618** can then compare the desired output (+V) with the actual output (**0**) and output −V. In this case, −V is equivalent to ground.

The op-amp outputs and grounds the pre-diode junction (i.e., see node A) and causes a greater electric field across inputs **1** and **4** and the layer **1** output. This increased electric field (larger voltage drop) can cause the nanoconductors in the solution between the electrode junctions to align themselves, aggregate, and form a stronger connection between the 1 and 4 electrodes. Feedback can continue to be applied until output of physical neural network **600** matches the desired output. The same procedure can be applied to every output.

In accordance with the aforementioned example, assume that Output **1** was higher than the desired output (i.e., desired output **622**). If this were the case, the op-amp output can be +V and the connection between inputs **1** and **4** and layer one output can be raised to +V. Columbic repulsions between the nanoconductors can force the connection apart, thereby weakening the connection. The feedback will then continue until the desired output is obtained. This is just one training mechanism. One can see that the training mechanism does not require any computations, because it is a simple feedback mechanism.

Such a training mechanism, however, may be implemented in many different forms. Basically, the connections in a connection network must be able to change in accordance with the feedback provided. In other words, the very general notion of connections being strengthened or connections being weakened in a physical system is the essence of a physical neural network (i.e., Knowm). Thus, it can be appreciated that the training of such a physical neural network may not require a "CPU" to calculate connection values thereof. The Knowm can adapt itself. Complicated neural network solutions could be implemented very rapidly "on the fly", much like a human brain adapts as it performs.

The physical neural network disclosed herein thus has a number of broad applications. The core concept of a Knowm, however, is basic. The very basic idea that the connection values between electrode junctions by nanoconductors can be used in a neural network devise is all that required to develop an enormous number of possible configurations and applications thereof.

Another important feature of a physical neural network is the ability to form negative connections. This is an important feature that makes possible inhibitory effects useful in data processing. The basic idea is that the presence of one input can inhibit the effect of another input. In artificial neural networks as they currently exist, this is accomplished by multiplying the input by a negative connection value. Unfortunately, with a Knowm-based device, the connection may only take on zero or positive values under such a scenario

In other words, either there can be a connection or no connection. A connection can simulate a negative connection by dedicating a particular connection to be negative, but one connection cannot begin positive and through a learning process change to a negative connection. In general, if starts positive, it can only go to zero. In essence, it is the idea of possessing a negative connection initially that results in the simulation, because this does not occur in a human brain. Only one type of signal travels through axon/dendrites in a human brain. That signal is transferred into the flow of a neurotransmitter whose effect on the postsynaptic neuron can be either excitatory or inhibitory, depending on the neuron.

One method for solving this problem is to utilize two sets of connections for the same output, having one set represent

the positive connections and the other set represent the negative connections. The output of these two layers can be compared, and the layer with the greater output will output either a high signal or a low signal, depending on the type of connection set (inhibitory or excitatory). This can be seen in FIG. **7**.

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network **700** that can be implemented in accordance with an alternative embodiment. Physical neural network **700** thus comprises a plurality of inputs **702** (not necessarily binary) which are respectively fed to layers **704**, **706**, **708**, and **710**. Each layer is analogous to the layers depicted earlier, such as for example layers **558** and **560** of FIG. **5**. An output **713** of layer **704** can be connected to a resistor **712**, a transistor **720** and a first input **727** of amplifier **726**. Transistor **720** is generally coupled between ground **701** and first input **727** of amplifier **726**. Resistor **712** is connected to a ground **701**. Note that ground **701** is analogous to ground **602** illustrated in FIG. **6** and ground **210** depicted in FIG. **2**. A second input **729** of amplifier **726** can be connected to a threshold voltage **756**. The output of amplifier **726** can in turn be fed to an inverting amplifier **736**.

The output of inverting amplifier **736** can then be input to a NOR device **740**. Similarly, an output **716** of layer **706** may be connected to resistor **714**, transistor **733** and a first input **733** of an amplifier **728**. A threshold voltage **760** is connected to a second input **737** of amplifier **728**. Resistor **714** is generally coupled between ground **701** and first input **733** of amplifier **728**. Note that first input **733** of amplifier **728** is also generally connected to an output **715** of layer **706**. The output of amplifier **728** can in turn be provided to NOR device **740**. The output from NOR device **740** is generally connected to a first input **745** of an amplifier **744**. An actual output **750** can be taken from first input **745** to amplifier **744**. A desired output **748** can be taken from a second input **747** to amplifier **744**. The output from amplifier **744** is generally provided at node A, which in turn is connected to the input to transistor **720** and the input to transistor **724**. Note that transistor **724** is generally coupled between ground **701** and first input **733** of amplifier **728**. The second input **731** of amplifier **728** can produce a threshold voltage **760**.

Layer **708** provides an output **717** that can be connected to resistor **716**, transistor **725** and a first input **737** to an amplifier **732**. Resistor **716** is generally coupled between ground **701** and the output **717** of layer **708**. The first input **737** of amplifier **732** is also electrically connected to the output **717** of layer **708**. A second input **735** to amplifier **732** may be tied to a threshold voltage **758**. The output from amplifier **732** can in turn be fed to an inverting amplifier **738**. The output from inverting amplifier **738** may in turn be provided to a NOR device **742**. Similarly, an output **718** from layer **710** can be connected to a resistor **719**, a transistor **728** and a first input **739** of an amplifier **734**. Note that resistor **719** is generally coupled between node **701** and the output **719** of layer **710**. A second input **741** of amplifier **734** may be coupled to a threshold voltage **762**. The output from of NOR device **742** is generally connected to a first input **749** of an amplifier **746**. A desired output **752** can be taken from a second input **751** of amplifier **746**. An actual output **754** can be taken from first input **749** of amplifier **746**. The output of amplifier **746** may be provided at node B, which in turn can be tied back to the respective inputs to transistors **725** and **728**. Note that transistor **725** is generally coupled between ground **701** and the first input **737** of amplifier **732**. Similarly, transistor **728** is generally connected between ground **701** and the first input **739** of amplifier **734**.

Note that transistors **720**, **724**, **725** and/or **728** each can essentially function as a switch to ground. A transistor such as, for example, transistor **720**, **724**, **725** and/or **728** may comprise a field-effect transistor (FET) or another type of transistor, such as, for example, a single-electron transistor (SET). Single-electron transistor (SET) circuits are essential for hybrid circuits combining quantum SET devices with conventional electronic devices. Thus, SET devices and circuits may be adapted for use with the physical neural network of the embodiments. This is particularly important because as circuit design rules begin to move into regions of the sub-100 nanometer scale, where circuit paths are only 0.001 of the thickness of a human hair, prior art device technologies will begin to fail, and current leakage in traditional transistors will become a problem. SET offers a solution at the quantum level, through the precise control of a small number of individual electrons. Transistors such as transistors **720**, **724**, **725** and/or **728** can also be implemented as carbon nanotube transistors.

A truth table for the output of circuit **700** is illustrated at block **780** in FIG. **7**. As indicated at block **780**, when an excitatory output is high and the inhibitory output is also high, the final output is low. When the excitatory output is high and the inhibitory output is low, the final output is high. Similarly, when the excitatory output is low and the inhibitory output is high, the final output is low. When the excitatory output is low and the inhibitory output is also low, the final output is low. Note that layers **704** and **708** may thus comprise excitatory connections, while layers **706** and **710** may comprise inhibitory connections.

For every desired output, two sets of connections are used. The output of a two-diode neuron can be fed into an op-amp (e.g., a comparator). If the output that the op-amp receives is low when it should be high, the op-amp outputs a low signal. This low signal can cause the transistors (e.g., transistors **720**, **725**) to saturate and ground out the pre-diode junction for the excitatory diode. Such a scenario can cause, as indicated previously, an increase in the voltage drop across those connections that need to increase their strength. Note that only those connections going to the excitatory diode are strengthened. Likewise, if the desired output were low when the actual output was high, the op-amp can output a high signal. This can cause the inhibitory transistor (e.g., an NPN transistor) to saturate and ground out the neuron junction of the inhibitory connections. Those connections going to the inhibitory diode can thereafter strengthen.

At all times during the learning process, a weak alternating electric field can be applied perpendicular to the connections. This can cause the connections to weaken by rotating the nanotube perpendicular to the connection direction. This perpendicular field is important because it can allow for a much higher degree of adaptation. To understand this, one must realize that the connections cannot (practically) keep getting stronger and stronger. By weakening those connections not contributing much to the desired output, we decrease the necessary strength of the needed connections and allow for more flexibility in continuous training. This perpendicular alternating voltage can be realized by the addition of two electrodes on the outer extremity of the connection set, such as plates sandwiching the connections (i.e., above and below). Other mechanisms, such as increasing the temperature of the nanotube suspension could also be used for such a purpose, although this method is perhaps a little less controllable or practical.

The circuit depicted in FIG. **7** can be separated into two separate circuits. The first part of the circuit can be composed of nanotube connections, while the second part of the circuit comprises the "neurons" and the learning mechanism (i.e., op-amps/comparator). The learning mechanism on first glance appears similar to a relatively standard circuit that could be implemented on silicon with current technology. Such a silicon implementation can thus comprise the "neuron" chip. The second part of the circuit (i.e., the connections) is thus a new type of chip, although it could be constructed with current technology. The connection chip can be composed of an orderly array of electrodes spaced anywhere from, for example, 100 nm to 1 μm or perhaps even further. In a biological system, one talks of synapses connecting neurons. It is in the synapses where the information is processed, (i.e., the "connection weights"). Similarly, such a chip can contain all of the synapses for the physical neural network. A possible arrangement thereof can be seen in FIG. **8**.

FIG. **8** illustrates a schematic diagram of a chip layout **800** for a connection network that may be implemented in accordance with an alternative embodiment. FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., connection network **800**) that can be implemented in accordance with one or more embodiments. Chip layout **800** includes an input array composed of plurality of inputs **801**, **802**, **803**, **804**, and **805**, which are provided to a plurality of layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. A plurality of outputs **802** can be derived from layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. Inputs **801** can be coupled to layers **806** and **807**, while inputs **802** can be connected to layers **808** and **809**. Similarly, inputs **803** can be connected to layers **810** and **811**. Also, inputs **804** can be connected to layers **812** and **813**. Inputs **805** are generally connected to layers **814** and **815**.

Similarly, such an input array can includes a plurality of inputs **831**, **832**, **833**, **834** and **835** which are respectively input to a plurality of layers **816**, **817**, **818**, **819**, **820**, **821**, **822**, **823**, **824** and **825**. Thus, inputs **831** can be connected to layers **816** and **817**, while inputs **832** are generally coupled to layers **818** and **819**. Additionally, inputs **833** can be connected to layers **820** and **821**. Inputs **834** can be connected to layers **822** and **823**. Finally, inputs **835** are connected to layers **824** and **825**. Arrows **828** and **830** represent a continuation of the aforementioned connection network pattern. Those skilled in the art can appreciate, of course, that chip layout **800** is not intended to represent an exhaustive chip layout or to limit the scope of the invention. Many modifications and variations to chip layout **800** are possible in light of the teachings herein without departing from the scope of the embodiments. It is contemplated that the use of a chip layout, such as chip layout **800**, can involve a variety of components having different characteristics.

Preliminary calculations based on a maximum etching capability of 200 nm resolution indicated that over 4 million synapses could fit on an area of approximately 1 cm^2. The smallest width that an electrode can possess is generally based on current lithography. Such a width may of course change as the lithographic arts advance. This value is actually about 70 nm for state-of-the-art techniques currently. These calculations are of course extremely conservative, and are not considered a limiting feature of the embodiments. Such calculations are based on an electrode with, separation, and gap of approximately 200 nm. For such a calculation, for example, 166 connection networks comprising 250 inputs and 100 outputs can fit within a one square centimeter area.

If such chips are stacked vertically, an untold number of synapses could be attained. This is two to three orders of magnitude greater than some of the most capable neural network chips out there today, chips that rely on standard methods to calculate synapse weights. Of course, the geometry of the chip could take on many different forms, and it is quite possible (based on a conservative lithography and chip layout) that many more synapses could fit in the same space. The training of a chip this size would take a fraction of the time of a comparably sized traditional chip using digital technology.

The training of such a chip is primarily based on two assumptions. First, the inherent parallelism of a physical neural network (i.e., a Knowm) can permit all training sessions to occur simultaneously, no matter now large the associated connection network. Second, recent research has indicated that near perfect aligning of nanotubes can be accomplished in approximately 15 minutes. If one considers that the input data, arranged as a vector of binary "high's" and "low's" is presented to the Knowm simultaneously, and that all training vectors are presented one after the other in rapid succession (e.g., perhaps 100 MHz or more), then each connection would "see" a different frequency in direct proportion to the amount of time that its connection is required for accurate data processing (i.e., provided by a feedback mechanism). Thus, if it only takes approximately 15 minutes to attain an almost perfect state of alignment, then this amount of time would comprise the longest amount of time required to train, assuming that all of the training vectors are presented during that particular time period.

FIG. **9** illustrates a flow chart **900** of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment. Initially, as indicated at block **902**, a connection gap is created from a connection network structures. As indicated earlier, the goal for such a connection network is generally to develop a network of connections of "just" the right values to satisfy particular information processing requirements, which is precisely what a neural network accomplishes. As illustrated at block **904**, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent," and also a suspension.

The solvent utilized can comprise a volatile liquid that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution may be sandwiched between wafers of silicon or other materials. If the fluid has a melting point that is approximately at room temperature, then the viscosity of the fluid could be controlled easily. Thus, if it is desired to lock the connection values into a particular state, the associated physical neural network (i.e., Knowm) may be cooled slightly until the fluid freezes. The term "solvent" as utilized herein thus can include fluids such as for example, toluene, hexadecane, mineral oil, etc. Note that the solution in which the nanoconductors (i.e., nanoconnections) are present should generally comprise a dielectric. Thus, when the resistance between the electrodes is measured, the conductivity of the nanoconductors can be essentially measured, not that of the solvent. The nanoconductors can be suspended in the solution or can alternately lie on the bottom surface of the connection gap. The solvent may also be provided in the form of a gas.

As illustrated thereafter at block **906**, the nanoconductors must be suspended in the solvent, either dissolved or in a suspension of sorts, but generally free to move around,

either in the solution or on the bottom surface of the gap. As depicted next at block **908**, the electrical conductance of the solution must be less than the electrical conductance of the suspended nanoconductor(s). Similarly, the electrical resistance of the solution is greater than the electrical resistance of the nanoconductor.

Next, as illustrated at block **910**, the viscosity of the substance should not be too much so that the nanoconductors cannot move when an electric field (e.g., voltage) is applied. Finally, as depicted at block **912**, the resulting solution of the "solvent" and the nanoconductors is thus located within the connection gap.

Note that although a logical series of steps is illustrated in FIG. **9**, it can be appreciated that the particular flow of steps can be re-arranged. Thus, for example, the creation of the connection gap, as illustrated at block **902**, may occur after the preparation of the solution of the solvent and nanoconductor(s), as indicated at block **904**. FIG. **9** thus represents merely possible series of steps, which may be followed to create a connection network. A variety of other steps may be followed as long as the goal of achieving a connection network is achieved. Similar reasoning also applies to FIG. **10**.

FIG. **10** illustrates a flow chart **1000** of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment. As indicated at block **1002**, an electric field can be applied across the connection gap discussed above with respect to FIG. **9**. The connection gap can be occupied by the solution discussed above. As indicated thereafter at block **1004**, to create the connection network, the input terminals can be selectively raised to a positive voltage while the output terminals are selectively grounded. As illustrated thereafter at block **1006**, connections thus form between the inputs and the outputs. The important requirements that make the resulting physical neural network functional as a neural network is that the longer this electric field is applied across the connection gap, or the greater the frequency or amplitude, the more nanoconductors align and the stronger the connection becomes. Thus, the connections that get utilized the most frequently become the strongest.

As indicated at block **1008**, the connections can either be initially formed and have random resistances or no connections will be formed at all. By forming initial random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up as much. Depending on the rate of connection decay, having initial random connections could prove to be a faster method, although not necessarily. A connection network can adapt itself to whatever is required regardless of the initial state of the connections. Thus, as indicated at block **1010**, as the electric field is applied across the connection gap, the more the nonconductor(s) will align and the stronger the connection becomes. Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block **1012**. As illustrated at block **1014**, the resistance of the connection can be maintained or lowered by selective activations of the connections. In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain.

The neurons in a human brain, although seemingly simple when viewed individually, interact in a complicated network that computes with both space and time. The most basic picture of a neuron, which is usually implemented in technology, is a summing device that adds up a signal. Actually, this statement can be made even more general by stating that a neuron adds up a signal in discrete units of time. In other words, every group of signals incident upon the neuron can be viewed as occurring in one moment in time. Summation thus occurs in a spatial manner. The only difference between one signal and another signal depends on where such signals originate. Unfortunately, this type of data processing excludes a large range of dynamic, varying situations that cannot necessarily be broken up into discrete units of time.

The example of speech recognition is a case in point. Speech occurs in the time domain. A word is understood as the temporal pronunciation of various syllables. A sentence is composed of the temporal separation of varying words. Thoughts are composed of the temporal separation of varying sentences. Thus, for an individual to understand a spoken language at all, a syllable, word, sentence or thought must exert some type of influence on another syllable, word, sentence or thought. The most natural way that one sentence can exert any influence on another sentence, in the light of neural networks, is by a form of temporal summation. That is, a neuron "remembers" the signals it received in the past.

The human brain accomplishes this feat in an almost trivial manner. When a signal reaches a neuron, the neuron has an influx of ions rush through its membrane. The influx of ions contributes to an overall increase in the electrical potential of the neuron. Activation is achieved when the potential inside the cell reaches a certain threshold. The one caveat is that it takes time for the cell to pump out the ions, something that it does at a more or less constant rate. So, if another signal arrives before the neuron has time to pump out all of the ions, the second signal will add with the remnants of the first signal and achieve a raised potential greater than that which could have occurred with only the second signal. The first signal influences the second signal, which results in temporal summation.

Implementing this in a technological manner has proved difficult in the past. Any simulation would have to include a "memory" for the neuron. In a digital representation, this requires data to be stored for every neuron, and this memory would have to be accessed continually. In a computer simulation, one must discritize the incoming data, since operations (such as summations and learning) occur serially. That is, a computer can only do one thing at a time. Transformations of a signal from the time domain into the spatial domain require that time be broken up into discrete lengths, something that is not necessarily possible with real-time analog signals in which no point exists within a time-varying signal that is uninfluenced by another point.

A physical neural network, however, is generally not digital. A physical neural network is a massively parallel analog device. The fact that actual molecules (e.g., nanoconductors) must move around (in time) makes temporal summation a natural occurrence. This temporal summation is built into the nanoconnections. The easiest way to understand this is to view the multiplicity of nanoconnections as one connection with one input into a neuron-like node (Op-amp, Comparator, etc.). This can be seen in FIG. **11**.

FIG. **11** illustrates a schematic diagram of a circuit **1100** illustrating temporal summation within a neuron, in accordance with a preferred embodiment. As indicated in FIG. **11**, an input **1102** can be provided to nanoconnections **1104**, which in turn can provide a signal, which can be input to an amplifier **1110** (e.g., op amp) at node B. A resistor **1106** can be connected to node A, which in turn is electrically equivalent to node B. Node B can be connected to a negative input of amplifier **1100**. Resistor **1108** can also be connected to a ground **1108**. Amplifier **1110** provides output **1114**. Note that although nanoconnections **1104** is referred to in the plural it

can be appreciated that nanoconnections **1104** can comprise a single nanoconnection or a plurality of nanoconnections. For simplicity sake, however, the plural form is used to refer to nanoconnections **1104**.

Input **1102** can be provided by another physical neural network (i.e., Knowm) to cause increased connection strength of nanoconnections **1104** over time. This input would most likely arrive in pulses, but could also be continuous. A constant or pulsed electric field perpendicular to the connections can serve to constantly erode the connections, so that only signals of a desired length or amplitude can cause a connection to form. Once the connection is formed, the voltage divider formed by nanoconnection **1104** and resistor **1106** can cause a voltage at node A in direct proportion to the strength of nanoconnections **1104**. When the voltage at node A reaches a desired threshold, the amplifier (i.e., an op-amp and/or comparator), will output a high voltage (i.e., output **1114**). The key to the temporal summation is that, just like a real neuron, it takes time for the electric field to breakdown the nanoconnections **1104**, so that signals arriving close in time will contribute to the firing of the neuron (i.e., op-amp, comparator, etc.). Temporal summation has thus been achieved. The parameters of the temporal summation could be adjusted by the amplitude and frequency of the input signals and the perpendicular electric field.

FIG. **12** illustrates a block diagram illustrating a pattern recognition system **1200**, which may be implemented with a physical neural network device **1222**, in accordance with an alternative embodiment. Note that pattern recognition system **1200** can be implemented as a speech recognition system. Although pattern recognition system **1200** is depicted herein in the context of speech recognition, a physical neural network device (i.e., a Knowm device) may be implemented with other pattern recognition systems, such as visual and/or imaging recognition systems. FIG. **12** thus does not comprise a limiting feature of the embodiments and is presented for general edification and illustrative purposes only. Those skilled in the art can appreciate that the diagram depicted in FIG. **12** may be modified as new applications and hardware are developed. The development or use of a pattern recognition system such as pattern recognition system **1200** of FIG. **12** by no means limits the scope of the physical neural network (i.e., Knowm) disclosed herein.

FIG. **12** thus illustrates in block diagram fashion, the system structure of a speech recognition device using a neural network according to an alternative embodiment. The pattern recognition system **1200** can be provided with a CPU **1211** for performing the functions of inputting vector rows and instructor signals (vector rows) to an output layer for the learning process of a physical neural network device **1222**, and changing connection weights between respective neuron devices based on the learning process. Pattern recognition system **1200** can be implemented within the context of a data-processing system, such as, for example, a personal computer or personal digital assistant (PDA), both of which are well known in the art.

The CPU **1211** can perform various processing and controlling functions, such as pattern recognition, including but not limited to speech and/or visual recognition based on the output signals from the physical neural network device **1222**. The CPU **1211** is connected to a read-only memory (ROM) **1213**, a random-access memory (RAM) **1214**, a communication control unit **1215**, a printer **1216**, a display unit **1217**, a keyboard **1218**, an FFT (fast Fourier transform) unit **1221**, a physical neural network device **1222** and a graphic reading unit **1224** through a bus line **1220** such as a data bus line. The bus line **1220** may comprise, for example, an ISA, EISA, or PCI bus.

The ROM **1213** is a read-only memory storing various programs or data used by the CPU **1211** for performing processing or controlling the learning process, and speech recognition of the physical neural network device **1222**. The ROM **1213** may store programs for carrying out the learning process according to error back-propagation for the physical neural network device or code rows concerning, for example, 80 kinds of phonemes for performing speech recognition. The code rows concerning the phonemes can be utilized as second instructor signals and for recognizing phonemes from output signals of the neuron device network. Also, the ROM **1213** can store programs of a transformation system for recognizing speech from recognized phonemes and transforming the recognized speech into a writing (i.e., written form) represented by characters.

A predetermined program stored in the ROM **1213** can be downloaded and stored in the RAM **1214**. RAM **1214** generally functions as a random access memory used as a working memory of the CPU **1211**. In the RAM **1214**, a vector row storing area can be provided for temporarily storing a power obtained at each point in time for each frequency of the speech signal analyzed by the FFT unit **1221**. A value of the power for each frequency serves as a vector row input to a first input portion of the physical neural network device **1222**. Further, in the case where characters or graphics are recognized in the physical neural network device, the image data read by the graphic reading unit **1224** are stored in the RAM **1214**.

The communication control unit **1215** transmits and/or receives various data such as recognized speech data to and/or from another communication control unit through a communication network **1202** such as a telephone line network, an ISDN line, a LAN, or a personal computer communication network. Network **1202** may also comprise, for example, a telecommunications network, such as a wireless communications network. Communication hardware methods and systems thereof are well known in the art.

The printer **1216** can be provided with a laser printer, a bubble-type printer, a dot matrix printer, or the like, and prints contents of input data or the recognized speech. The display unit **1217** includes an image display portion such as a CRT display or a liquid crystal display, and a display control portion. The display unit **1217** can display the contents of the input data or the recognized speech as well as a direction of an operation required for speech recognition utilizing a graphical user interface (GUI).

The keyboard **1218** generally functions as an input unit for varying operating parameters or inputting setting conditions of the FFT unit **1221**, or for inputting sentences. The keyboard **1218** is generally provided with a ten-key numeric pad for inputting numerical figures, character keys for inputting characters, and function keys for performing various functions. A mouse **1219** can be connected to the keyboard **1218** and serves as a pointing device.

A speech input unit **1223**, such as a microphone can be connected to the FFT unit **1221**. The FFT unit **1221** transforms analog speech data input from the voice input unit **1223** into digital data and carries out spectral analysis of the digital data by discrete Fourier transformation. By performing a spectral analysis using the FFT unit **1221**, the vector row based on the powers of the respective frequencies are output at predetermined intervals of time. The FFT unit **1221** performs an analysis of time-series vector rows, which represent characteristics of the inputted speech. The vector

rows output by the FFT **1221** are stored in the vector row storing area in the RAM **1214**.

The graphic reading unit **224**, provided with devices such as a CCD (Charged Coupled Device), can be used for reading images such as characters or graphics recorded on paper or the like. The image data read by the image-reading unit **1224** are stored in the RAM **1214**. Note that an example of a pattern recognition apparatus, which may be modified for use with the physical neural network described herein, is disclosed in U.S. Pat. No. 6,026,358 to Tomabechi, Feb. 16, 2000, "Neural Network, A Method of Learning of a Neural Network and Phoneme Recognition Apparatus Utilizing a Neural Network." U.S. Pat. No. 6,026,358 is incorporated herein by reference. It can be appreciated that the Tomabechi reference does not teach, suggest or anticipate the embodiments, but is discussed herein for general illustrative, background and general edification purposes only.

The implications of a physical neural network are tremendous. With existing lithography technology, many electrodes in an array such as depicted in FIG. **5** can be etched onto a wafer of silicon. The neuron-diodes, as well as the training circuitry illustrated in FIG. **6**, could be built onto the same silicon wafer, although it may be desirable to have the connections on a separate chip due to the liquid solution of nanoconductors. A solution of suspended nanoconductors could be placed between the electrode connections and the chip could be packaged. The resulting "chip" would look much like a current Integrated Chip (IC) or VLSI (very large scale integrated) chips. One could also place a rather large network parallel with a computer processor as part of a larger system. Such a network, or group of networks, could add significant computational capabilities to standard computers and associated interfaces.

For example, such a chip may be constructed utilizing a standard computer processor in parallel with a large physical neural network or group of physical neural networks. A program can then be written such that the standard computer teaches the neural network to read, or create an association between words, which is precisely the same sort of task in which neural networks can be implemented. Once the physical neural network is able to read, it can be taught for example to "surf" the Internet and find material of any particular nature. A search engine can then be developed that does not search the Internet by "keywords", but instead by meaning. This idea of an intelligent search engine has already been proposed for standard neural networks, but until now has been impractical because the network required was too big for a standard computer to simulate. The use of a physical neural network (i.e., physical neural network) as disclosed herein now makes a truly intelligent search engine possible.

A physical neural network can be utilized in other applications, such as, for example, speech recognition and synthesis, visual and image identification, management of distributed systems, self-driving cars, filtering, etc. Such applications have to some extent already been accomplished with standard neural networks, but are generally limited in expense, practicality and not very adaptable once implemented. The use of a physical neural network can permit such applications to become more powerful and adaptable. Indeed, anything that requires a bit more "intelligence" could incorporate a physical neural network. One of the primary advantages of a physical neural network is that such a device and applications thereof can be very inexpensive to manufacture, even with present technology. The lithographic techniques required for fabricating the electrodes and channels therebetween has already been perfected and implemented in industry.

Most problems in which a neural network solution is implemented are complex adaptive problems, which change in time. An example is weather prediction. The usefulness of a physical neural network is that it could handle the enormous network needed for such computations and adapt itself in real-time. An example wherein a physical neural network (i.e., Knowm) can be particularly useful is the Personal Digital Assistant (PDA). PDA's are well known in the art. A physical neural network applied to a PDA device can be advantageous because the physical neural network can ideally function with a large network that could constantly adapt itself to the individual user without devouring too much computational time from the PDA. A physical neural network could also be implemented in many industrial applications, such as developing a real-time systems control to the manufacture of various components. This systems control can be adaptable and totally tailored to the particular application, as necessarily it must.

It will be appreciated that variations of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also that various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

What is claimed is:

1. A utilized nanotechnology apparatus for maintaining components in a neural network, said apparatus comprising:
- a connection gap between two terminals;
- a solution having a melting point at an approximately room temperature, said solution is maintained in said connection gap, said solution having a plurality of nanoparticles forming nanoconnections, said nanoparticles having connection strengths;
- a neural network in use with said solution and said connection gap to utilize nanotechnology; and
- a utilized nanotechnology system for locking into place said connection strengths as a result of freezing said solution when power is removed from said neural network.

2. The apparatus of claim **1** wherein said solution comprises a solvent.

3. The apparatus of claim **2** wherein said plurality of nanoparticles are suspended in said solvent.

4. The apparatus of claim **2** wherein said solvent comprises a gas.

5. The apparatus of claim **2** wherein said solvent comprises a liquid.

6. The apparatus of claim **2** wherein said solvent comprises a fluid.

7. The apparatus of claim **1** wherein said plurality of nanoparticles comprise nanoconductors.

8. The apparatus of claim **7** wherein said solution comprises an electrical conductance that is less than that an electrical conductance of said nanoconductors.

9. The apparatus of **1** wherein said solution is sealed, thereby preventing exposure of said solution to air.

10. The apparatus of claim **1** wherein said solution comprises a dielectric.

11. A utilized nanotechnology apparatus for maintaining components in neural network, said apparatus comprising:
- a connection gap between two terminals;

a solution comprising a solvent having a melting point at an approximately room temperature, said solution is maintained in said connection gap, said solution having a plurality of nanoparticles suspended in said solvent and forming nanoconnections, said nanoparticles having having connection strengths;

a neural network in use with said solution and said connection gap to utilize nanotechnology; and

a utilized nanotechnology system for locking into place said connection strengths as a result of freezing said solution when power is removed from said neural network.

12. The apparatus of claim **11** wherein said plurality of nanoparticles comprise nanoconductors.

13. The apparatus of claim **12** wherein said solution comprises an electrical conductance that is less than that an electrical conductance of said nanoconductors.

14. The apparatus of **11** wherein said solution is sealed, thereby preventing exposure of said solution to air.

15. The apparatus of claim **11** wherein said solution comprises a dielectric.

* * * * *

US007028017B2

(12) **United States Patent**
Nugent

(10) **Patent No.: US 7,028,017 B2**
(45) **Date of Patent: Apr. 11, 2006**

(54) **TEMPORAL SUMMATION DEVICE UTILIZING NANOTECHNOLOGY**

(75) Inventor: **Alex Nugent**, Santa Fe, NM (US)

(73) Assignee: **Knowm Tech, LLC**, Albuquerque, NM (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 72 days.

(21) Appl. No.: **11/045,536**

(22) Filed: **Jan. 31, 2005**

(65) **Prior Publication Data**

US 2005/0149465 A1 Jul. 7, 2005

Related U.S. Application Data

(63) Continuation of application No. 10/095,273, filed on Mar. 12, 2002, now Pat. No. 6,889,216.

(51) **Int. Cl.**
G06F 15/18 (2006.01)
G06E 1/00 (2006.01)

(52) **U.S. Cl.** **706/26**; 706/30

(58) **Field of Classification Search** 706/26, 706/30

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,802,951	A	2/1989	Clark et al.	156/630
4,926,064	A *	5/1990	Tapang	706/26
4,974,146	A	11/1990	Works et al.	364/200
4,988,891	A	1/1991	Mashiko	307/201
5,315,162	A	5/1994	McHardy et al.	307/201
5,422,983	A	6/1995	Castelaz et al.	395/24
5,475,794	A	12/1995	Mashiko	395/24
5,589,692	A	12/1996	Reed	257/23
5,649,063	A	7/1997	Bose	395/22
5,706,404	A	1/1998	Colak	395/24
5,717,832	A	2/1998	Steimle et al.	395/24
5,783,840	A	7/1998	Randall et al.	257/24
5,812,993	A	9/1998	Ginosar et al.	706/26
5,904,545	A	5/1999	Smith et al.	438/455
5,951,881	A	9/1999	Rogers et al.	216/41
5,978,782	A	11/1999	Neely	706/16
6,026,358	A	2/2000	Tomabechi	704/232
6,128,214	A	10/2000	Kuekes et al.	365/151
6,245,630	B1	6/2001	Ishikawa	438/393
6,248,529	B1	6/2001	Connolly	435/6

(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 022 764 A1 1/2000

(Continued)

OTHER PUBLICATIONS

Peter Weiss, "Circuitry in a Nanowire: Novel Growth Method May Transform Chips," Science News Online, vol. 161, No. 6; Feb. 9, 2002.

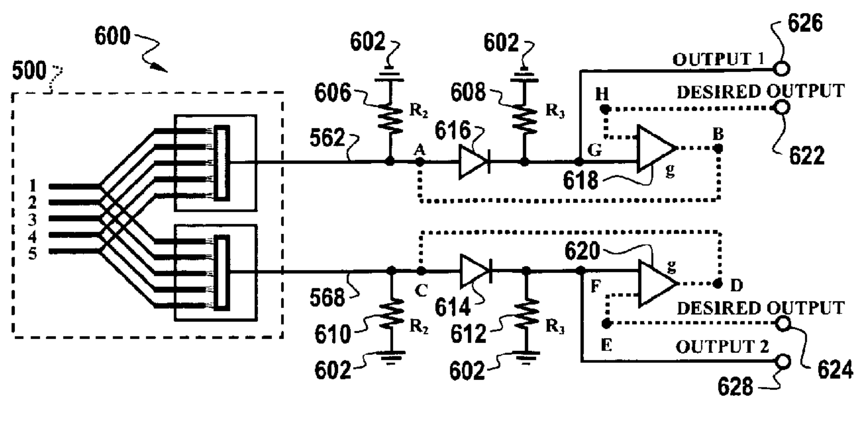
(Continued)

Primary Examiner—George Davis

(57) **ABSTRACT**

A temporal summation device can be composed of one or more nanoconnections having an input and an output thereof, wherein an input signal provided to the input causes one or more of the nanoconnection to experience an increase in connection strength thereof over time. Additionally, a voltage divider is formed by the nanoconnection(s) and a resistor connected to the output of the nanoconnection(s), such that the voltage divider provides a voltage at the output the nanoonnection(s) that is in direct proportion to the connection strength of nanoconnection(s). An amplifier is also connected to the voltage divider, wherein when the voltage provided by the voltage divider attains a desired threshold voltage, the amplifier attains a high voltage output thereby providing a temporal summation device thereof.

20 Claims, 6 Drawing Sheets



U.S. PATENT DOCUMENTS

6,256,767 B1	7/2001	Kuekes et al.	716/9
6,282,530 B1	8/2001	Huang	706/41
6,294,450 B1	9/2001	Chen et al.	438/597
6,314,019 B1	11/2001	Kuekes et al.	365/151
6,330,553 B1	12/2001	Uchikawa et al.	706/2
6,335,291 B1	1/2002	Freeman	438/706
6,339,227 B1	1/2002	Ellenbogen	257/40
6,359,288 B1	3/2002	Ying et al.	257/14
6,363,369 B1	3/2002	Liaw et al.	706/15
6,383,923 B1	5/2002	Brown et al.	438/666
6,389,404 B1	5/2002	Carson et al.	706/18
6,407,443 B1	6/2002	Chen et al.	257/616
6,418,423 B1	7/2002	Kambhatla et al.	706/15
6,420,092 B1	7/2002	Yang et al.	430/311
6,422,450 B1	7/2002	Zhou et al.	228/121.85
6,423,583 B1	7/2002	Avouris et al.	438/132
6,424,961 B1	7/2002	Ayala	706/25
6,426,134 B1	7/2002	Lavin et al.	428/300.1
2001/0004471 A1	6/2001	Zhang	427/372.2
2001/0023986 A1	9/2001	Mancevski	257/741
2001/0024633 A1	9/2001	Lee et al.	423/447.3
2001/0044114 A1	11/2001	Connolly	435/6
2002/0001905 A1	1/2002	Choi et al.	438/268
2002/0004136 A1	1/2002	Gao et al.	428/367
2002/0030205 A1	3/2002	Varshavsky	257/208
2002/0090468 A1	7/2002	Goto et al.	427/580
2002/0102353 A1	8/2002	Mauthner et al.	427/255.28
2003/0177450 A1	9/2003	Nugent	716/1
2003/0236760 A1	12/2003	Nugent	706/26
2004/0039717 A1	2/2004	Nugent	706/27
2004/0150010 A1	8/2004	Snider	257/209
2004/0153426 A1	8/2004	Nugent	706/25
2004/0162796 A1	8/2004	Nugent	706/27
2004/0193558 A1	9/2004	Nugent	706/25

FOREIGN PATENT DOCUMENTS

EP	1 046 613 A2	4/2000
EP	1 100 106 A2	5/2001
EP	1 069 206 A2	7/2001
EP	1 115 135 A1	7/2001
EP	1 134 304 A2	9/2001
WO	WO 00/44094	7/2000

OTHER PUBLICATIONS

Press Release, "Nanowire-based electronics and optics comes one step closer," Eureka Alert, American Chemical Society; Feb. 1, 2002.

Weeks et al., "High-pressure nanolithography using low-energy electrons from a scanning tunneling microscope," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 38-42; Dec. 12, 2001.

CMP Cientifica, "Nanotech: the tiny revolution"; CMP Cientifica, Nov. 2001.

Diehl, et al., "Self-Assembled, Deterministic Carbon Nanotube Wiring Networks," Angew. Chem. Int. Ed. 2002, 41, No. 2; Received Oct. 22, 2001.

G. Pirio, et al., "Fabrication and electrical characteristics of carbon nanotube field emission microcathodes with an integrated gate electrode," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 1-4, Oct. 2, 2001.

Leslie Smith, "An Introduction to Neural Networks," Center for Cognitive and Computational Neuroscience, Dept. of Computing & Mathematics, University of Stirling, Oct. 25, 1996; http//www.cs.stir.ac.uk/~Iss/NNIntro/InvSlides.html.

V. Derycke et al., "Carbon Nanotube Inter- and Intramolecular Logic Gates," American Chemical Society, Nano Letters, XXXX, vol. 0, No. 0, A-D.

Mark K. Anderson, "Mega Steps Toward the Nanochip," Wired News, Apr. 27, 2001.

Collins et al., "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," Science, vol. 292, pp. 706-709, Apr. 27, 2001.

Landman et al., "Metal-Semiconductor Nanocontacts: Silicon Nanowires," Physical Review Letters, vol. 85, No. 9, Aug. 28, 2000.

John G. Spooner, "Tiny tubes mean big chip advances," Cnet News.com, Tech News First, Apr. 26, 2001.

Jeong-Mi Moon et al., "High-Yield Purification Process of Singlewalled Carbon Nanotubes," J. Phys. Chem. B 2001, 105, pp. 5677-5681.

"A New Class of Nanostructure: Semiconducting Nanobelts Offer Potential for Nanosensors and Nanoelectronics," Mar. 12, 2001, http://www.sciencedaily.com/releases/2001/03/010309080953.htm.

Hermanson et al., "Dielectrophoretic Assembly of Electrically Functional Microwaves from Nanoparticle Suspensions," Materials Science, vol. 294, No. 5544, Issue of Nov. 2, 2001, pp. 1082-1086.

Press Release, "Toshiba Demonstrates Operation of Single-Electron Transistor Circuit at Room Temperature," Toshiba, Jan. 10, 2001.

J. Appenzeller et al., "Optimized contact configuration for the study of transport phenomena in ropes of single-wall carbon nanotubes," Applied Physics Letters, vol. 78, No. 21, pp. 3313-3315, May 21, 2001.

David Rotman, "Molecular Memory, Replacing silicon with organic molecules could mean tiny supercomputers," Technology Review, May 2001, p. 46.

Westervelt et al., "Molecular Electronics," NSF Functional Nanostructures Grant 9871810, NSF Partnership in Nanotechnology Conference, Jan. 29-30, 2001; http://www.unix.oit.umass.edu/~nano/NewFiles/FN19_Harvard.pdf.

Niyogi et al., "Chromatographic Purification of Soluble Single-Walled Carbon Nanotubes (s-SWNTs)," J. Am. Chem. Soc 2001, 123, pp. 733-734, Received Jul. 10, 2000.

Duan et al., "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices," Nature, vol. 409, Jan. 4, 2001, pp. 67-69.

Paulson, et al., "Tunable Resistance of a Carbon Nanotube-Graphite Interface," Science, vol. 290, Dec. 1, 2000, pp. 1742-1744.

Wei et al., "Reliability and current carrying capacity of carbon nanotubes," Applied Physics Letters, vol. 79, No. 8, Aug. 20, 2001, pp. 1172-1174.

Collins et al., "Nanotubes for Electronics," Scientific American, Dec. 2000, pp. 62-69.

Avouris et al., "Carbon nanotubes: nanomechanics, manipulation, and electronic devices," Applied Surface Science 141 (1999), pp. 201-209.

Smith et al., "Electric-field assisted assembly and alignment of metallic nanowires," Applied Physics Letters, vol. 77, No. 9, Aug. 28, 2000, pp. 1399-1401.

Hone et al., "Electrical and thermal transport properties of magnetically aligned single wall carbon nanotubes films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 666-668.

Smith et al., "Structural anisotropy of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 663-665.

Andriotis et al., "Various bonding configurations of transition-metal atoms on carbon nanotubes: Their effect on contact resistance," Applied Physics Letters, vol. 76, No. 26, Jun. 26, 2000, pp. 3890-3892.
Chen et al.,"Aligning single-wall carbon nanotubes with an alternating-current electric field," Applied Physics Letters, vol. 78, No. 23, Jun. 4, 2001, pp. 3714-3716.
Bezryadin et al., "Self-assembled chains of graphitized carbon nanoparticles," Applied Physics Letters, vol. 74, No. 18, May 3, 1999, pp. 2699-2701.
Bezryadin et al., "Evolution of avalanche conducting states in electrorheological liquids," Physical Review E, vol. 59, No. 6, Jun. 1999, pp. 6896-6901.
Liu et al., "Fullerene Pipes," Science, vol. 280, May 22, 1998, pp. 1253-1255.
Yamamoto et al., "Orientation and purification of carbon nanotubes using ac electrophoresis," J. Phys. D: Appl. Phys 31 (1998) L34-L36.
Bandow et al., "Purification of Single-Wall Carbon Nanotubes by Microfiltration," J. Phys. Chem. B 1997, 101, pp. 8839-8842.
Tohji et al.,"Purifying single walled nanotubes," Nature, vol. 383, Oct. 24, 1996, p. 679.
Dejan Rakovic, "Hierarchical Neural Networks and Brainwaves: Towards a Theory of Consciousness," Brain & Consciousness: Proc. ECPD Workshop (ECPD, Belgrade, 1997), pp. 189-204.
Dave Anderson & George McNeill, "Artificial Neural Networks Technology," A DACS (Data & Analysis Center for Software) State-of-the-Art Report, Contract No. F30602-89-C-0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, Aug. 20, 1992.
Greg Mitchell, "Sub-50 nm Device Fabrication Strategies," Project No. 890-00, Cornell Nanofabrication Facility, Electronics—p. 90-91, National Nanofabrication Users Network.
John-William DeClaris, "An Introduction to Neural Networks," http://www.ee.umd.edu/medlab/neural/nn1.html.
"Neural Networks," StatSoft, Inc., http://www.statsoftinc.com/textbook/stevnet.html.
Stephen Jones, "Neural Networks and the Computation Brain or Maters relating to Artificial Intelligence," The Brain Project, http://www.culture.com.au/brain_proj/neur_net.htm.
David W. Clark, "An Introduction to Neural Networks"; http://members.home.net/neuralnet/introtonn/index.htm.
"A Basic Introduction to Neural Networks"; http://blizzard.gis.uiuc.edu/htmldocs/Neural/neural.html.
Meyer et al., "Computational neural networks: a general purpose tool for nanotechnology," Abstract, 5th Foresight Conference on Molecular Nanotechnology; http://www.foresight.org/Conferences/MNT05/Abstracts/Meyeabst.html.
Saito et al., "A 1M Synapse Self-Learning Digital Neural Network Chip," ISSCC, pp. 6.5-1 to 6.5-10, IEEE 1998.
Espejo, et al, "A 16×16 Cellular Neural Network Chip for Connected Component Detection," Jun. 30, 1999; http://www.imsc.cnm.csic.es/Chipcat/espejo/chip-2.pdf.
Pati et al., "Neural Networks for Tactile Perception," Systems Research Center and Dept. of Electrical Engineering, University of Maryland and U.S Navel Research Laboratory. 1987.
http://www.isr.umd.edu/TechReports/ISR/1987/TR_87-123/TR_87-123.phtml.
Osamu Fujita, "Statistical estimation of the number of hidden units for feedforward neural networks," Neural Networks 11 (1998), pp. 851-859.
Abraham Harte, "Liquid Crystals Allow Large-Scale Alignment of Carbon Nanotubes," CURJ (Caltech Undergraduate Research Journal), Nov. 2001, vol. 1, No. 2, pp. 44-49.
"Quantum-Dot Arrays for Computation," ORNL Review vol. 34, No. 2, 2001, pp. 1-5 http://www.ornlgov/ORNLReview/v34_2_01/arrays.htm.
Jabri, M.A. et al., "Adaptive Analog VLSI Neural Systems," Chapman & Hall, London SE1 8HN, UK, 1996, pp. 92-95.
Lipson et al., "Automatic Design and Manufacture of Robotic Lifeforms," NATURE, vol. 406, Aug. 31, 2000, pp. 974-978.
Kunitoshi Yamamoto, et al., "Rapid Communication Orientation and Purification of Carbon Nanotubes Using AC Electrophoresis", J. Phys. D. Appl. Phys 31 (1998) L34-L36.
E.S. Snow, et al., "Random networks of carbon nanotubes as electronic material", Applied Physics Letters, vol. 82, No. 12, Mar. 31, 2003, pp. 2145-2147.
R. Martel, et al., "Ambipolar Electrical Transport in Semiconducting Single-Wall Carbon Nanotubes," Physical Review Letters, vol. 87, No. 25, Dec. 17, 2001, pp. 256805-1 to 256805-4.
S. Heinze, et al., "Carbon Nanotubes as Schottky Barrier Transistors", vol. 89, No. 10, Sep. 2, 2002, pp. 106801-1 to 106801-4.
M. Dubson, et al., "Measurement of the conductivity exponent in two-dimensional percolating networks: square lattice versus random-void continuum", Physical Review B, vol. 32, No. 11, Dec. 1, 1985, pp. 7621-7623.
D.J. Frank, et al., "Highly efficient algorithm for percolative transport studies in two dimensions", Physical Review B, vol. 37, No. 1, Jan. 1, 1988, pp. 302-307.
Uma R. Karmarkar, et al., "Mechanisms and signficance of spike-timing dependent plasticity," Biol. Cybern. 87, 373-382 (2002), Jan. 28, 2002.
Uma R. Karmarkar, et al., "A Model of Spike-Timing Dependent Plasticity: One or Two Coincidence Detectors?", J. Neurophysiol, vol. 88, pp. 507-513, Jul. 2002.
M.C.W. van Rossum, et al., "Stable Hebbian Learning from Spkke-Timing-Dependent Plasticity", The Journal of Neuroscience, Dec. 1, 2003, 20(23), pp. 8812-8821.
Xiaohui Xie, et al., "Spike-based learning rules and stabilization of persistent neural activity,".
Nace L. Golding, et al., "Dendritic spikes as a mechanism for cooperative long-term potentiation", NATURE, vol. 418, Jul. 18, 2002, pp. 326-330.
Ozgur Turel, et al., "Possible nanoelectronic implementation of neuromorphic networks", Dept. o f Physics and Astronomy, Stony Brook University.
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3; Sept. 9, 2003; American Chemical Society, pp. 1379-1382.
J.M. Tour, et al., "NanoCell Electronic Memories," J.Am. Chem.Soc. 2003, 125, pp. 13279-13283.
J. Zaumseil, et al., "Three-Dimensional and Multilayer Nanostructures Formed by Nanotransfer Printing," Nano Letters, 2003, vol. 3, No. 9; Jul. 31, 2003, American Chemical Society, pp. 1223-1227.
Charles D. Schaper, "Patterned Transfer of Metallic Thin Film Nanostructures by Water-Soluble Polymer Templates," Nano Letters, 2003, vol. 3, No. 9; Jul. 26, 2003, American Chemical Society, pp. 1305-1309.

C.A. Dyke, et al., "Unbundled and Highly Functionalized Carbon Nanotubes from Aqueous Reactions," Nano Letters, 2003, vol. 3, No. 9; Aug. 19, 2003, American Chemical Society, pp. 1215-1218.
J. Chung, et al., "Nanoscale Gap Fabrication by Carbon Nanotube-Extracted Lithography (CEL)", Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003, American Chemical Society, pp. 1029-1031.
O. Harnack, et al., "Rectifying Behavior of Electrically Aligned ZnO Nanorods," Nano Letters, 2003, vol. 3, No. 8; Jun. 24, 2003, American Chemical Society, pp. 1097-1101.
M. S. Kumar, et al., "Influence of electric field type on the assembly of single walled carbon nanotubes," Chemical Physics Letters 383 (2004), Dec. 2, 2003; pp. 235-239.
S.W. Lee, et al., "Dielectrophoresis and electrohydrodynamics-mediated fluidic assembly of silicon resistors," Applied Physics Letters, vol. 83, No. 18, Nov. 3, 2003, pp. 3833-3835.
R. Krupke, et al., "Simultaneous Deposition of Metallic Bundles of Single-walled Carbon Nanotubes Using Ac-dielectrophoresis," Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003; American Chemical Society, pp. 1019-1023.
K. Bradley, et al., "Flexible Nanotube Electronics," Nano Letters, 2003, vol. 3, No. 10; Aug. 9, 2003, America Chemical Society, pp. 1353-1355.
T.B. Jones, "Frequency-dependent orientation of isolated particle chains," Journal of Electrostatics, 25 (1990), Elsevier Science Publishers, pp. 231-244.
L.A. Nagahara, "Directed placement of suspended carbon nanotubes for nanometer-scale assembly," Applied Physics Letters, vol. 80, No. 20; May 20, 2003; pp. 3826-3828.
A. Bezryadin, et al., "Electrostatic trapping of single conducting nanoparticles between electrodes," Applied Physics Letters, 71 (9), Sep. 1, 1997, pp. 1273-1275.
S. Suzuki, et al., "Quantitative Analysis of DNA Orientation in Stationary AC Electric Fields Using Fluorescence Anisotropy," IEEE Transactions of Industry Applications, vol. 34, No. 1; Jan./Feb. 1998, pp. 75-83.
Phaedon Avouris, "Molecular Electronics with Carbon Nanotubes," Accounts of Chemical Research, vol. 35, No. 12; Jul. 31, 2002, pp. 1025-1034.
X. Liu, et al., "Electric-Field Induced Accumulation and Alignment of Carbon Nanotubes," 2002 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp. 31-34.
R. Krupke, et al., "Contacting single bundles of carbon nanotubes with alternating electric fields," Appl. Phys. A. 76, Oct. 28, 2002, pp. 397-400.
M. Law, et al., "Photochemical Sensing of NO_2 with SnO^2 Nanoribbon Nanosensors at Room Temperature," Angew. Chem. 2002, 114, Nr. 13, pp. 2511-2514.
J. Tour, et al., "Nanocell Logic Gates for Molecular Computing," IEEE Transactions on Nanotechnology, vol. 1, No. 2, Jun. 2002, pp. 100-109.
A. Leonardi, et al., "Simulation methodology for dielectrophoresis in microelectronic Lab-on-a-chip," Modeling and Simulation of Microsystems 2002, pp. 96-99.
J. Chung, et al., "Nanoscale Gap Fabrication and Integration of Carbon Nanotubes by Micromachining," Solid-State Sensor, Actuator and Microsystems Workshop, Jun. 2-6, 2003; Hilton Head Island, South Carolina, pp. 161-164.
L. Zheng, et al., "Towards Single Molecule Manipulation with Dielectrophoresis Using Nanoelectrodes," IEEE-NANO 2003, Aug. 12-14, 2003, Moscone Convention Center, San Francisco, CA; pp. 437-440, http://ieeenano2003.arc.nasa.gov/program_contents.pdf.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.
V.C. Moore, et al., "Indvidually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3, No. 10; American Chemical Society; Sep. 8, 2003; pp. 1379-1382.
R. Krupke, "Separation of Metallic from Semiconducting Single-Walled Carbon Nanotubes," Science, vol. 301; Jul. 18, 2003; pp. 344-347.
Wolfgang Maass, "On the Relevance of Time in Neural Computation and Learning," In M. Li and A. Maruoka, editors, *Proc. of the 8th International Conference on Algorithmic Learning Theory in Sendai (Japan)*, vol. 1316 of *Lecture Notes in Computer Science*, pp. 364-388. Springer (Berlin), 1997.
Wolfgang Maass, "Noisy Spiking Neurons with Temporal Coding have more Computational Power than Sigmoidal Neurons," In M. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems*, vol. 9, pp. 211-217. MIT Press (Cambridge), 1997. (pp. 1-13, including Appendix).
L. Perrinet, et al., "Emergence of filters from natural scenes in a sparse spike coding scheme," Neurocomputing, 2003, pp. 1-14, http://www.laurent.perrinet.free.fr/publi/perrinet03neurocomputing.pdf.
L. Perrinet, et al., "Coherence detection in a spiking neuron via Hebbian learning," Neurocomputing, 2002, vol. 44-46, No. C., pp. 817-822, http://www.laurent.perrinet.free.fr/publi/perrinet02.pdf.
A. Jarosz, et al., "An Introductory Note on Gaussian Correlated Random Matrix," Feb. 21, 2003, pp. 1-20 http://www.if.uj.edu.pl/pl/koloSMP/prace/rndmatrix.pdf.
K. Bradley, et al., "Influence of Mobile Ions on Nanotube Based FET Devices," Nano Letters, 2003, vol. 3, No. 5; American Chemical Society, Apr. 4, 2003; pp. 639-641.

* cited by examiner

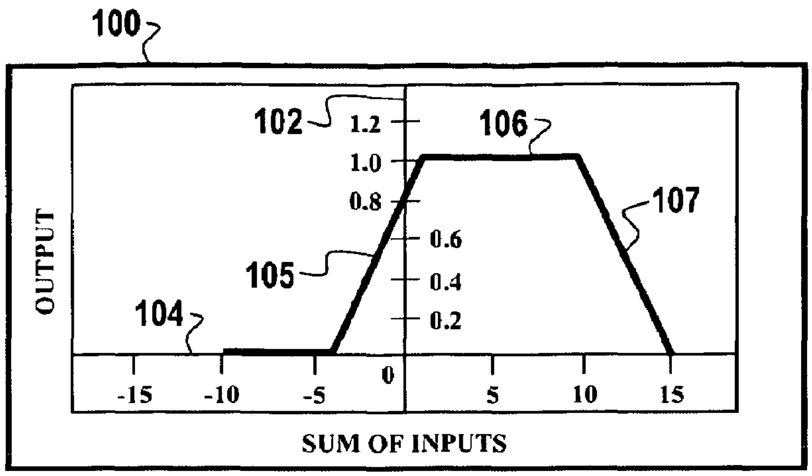


Figure 1

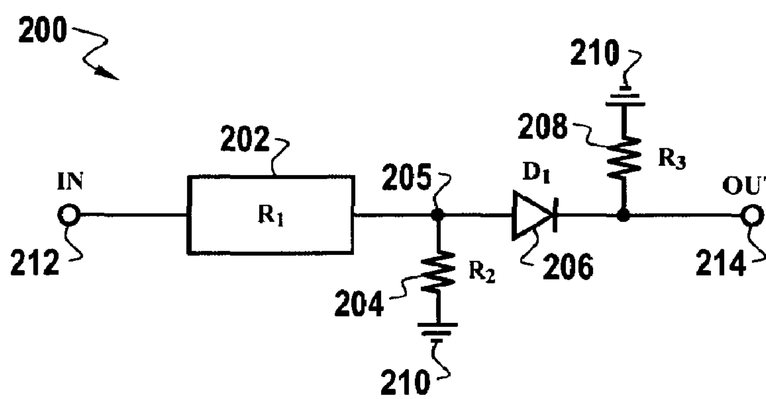


Figure 2

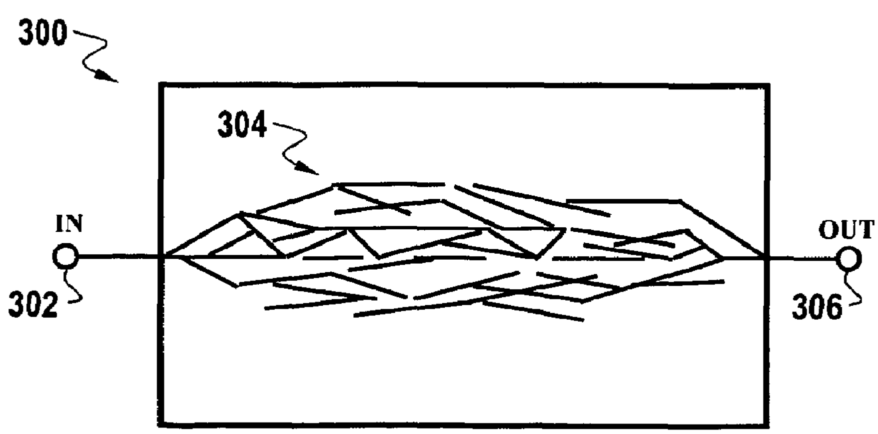


Figure 3

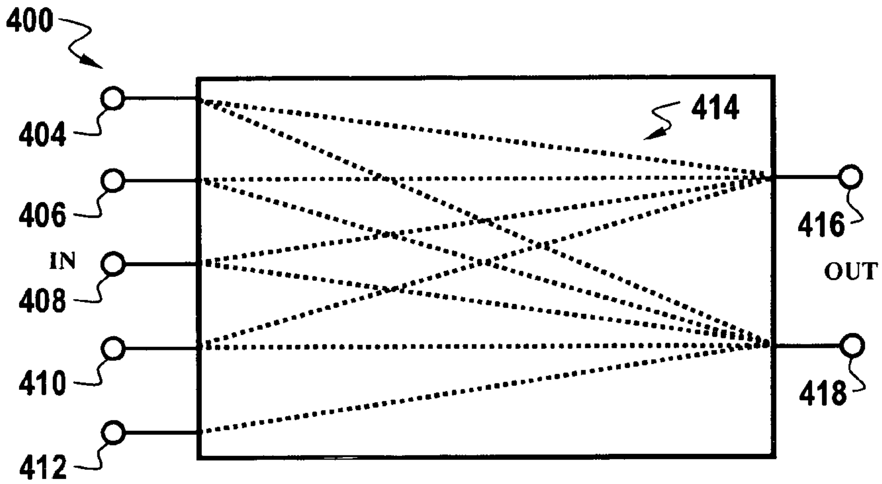


Figure 4

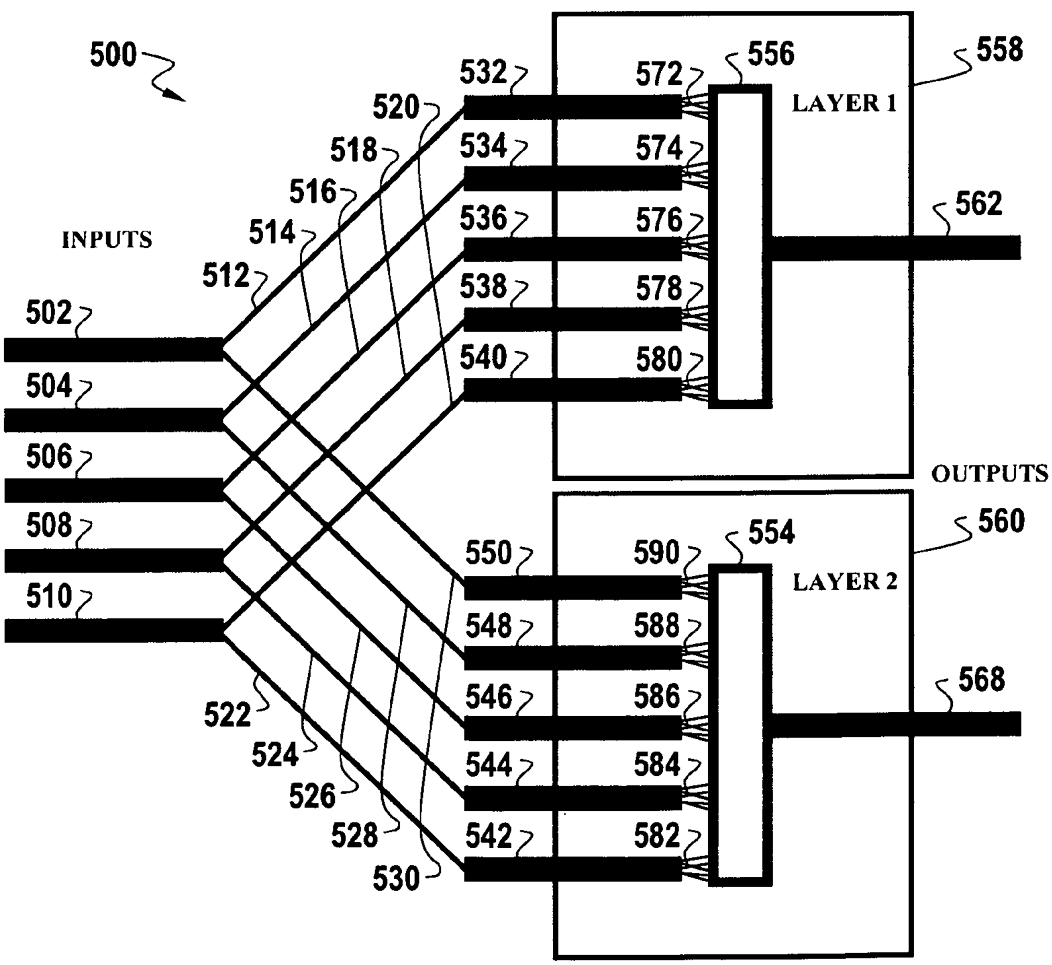


Figure 5

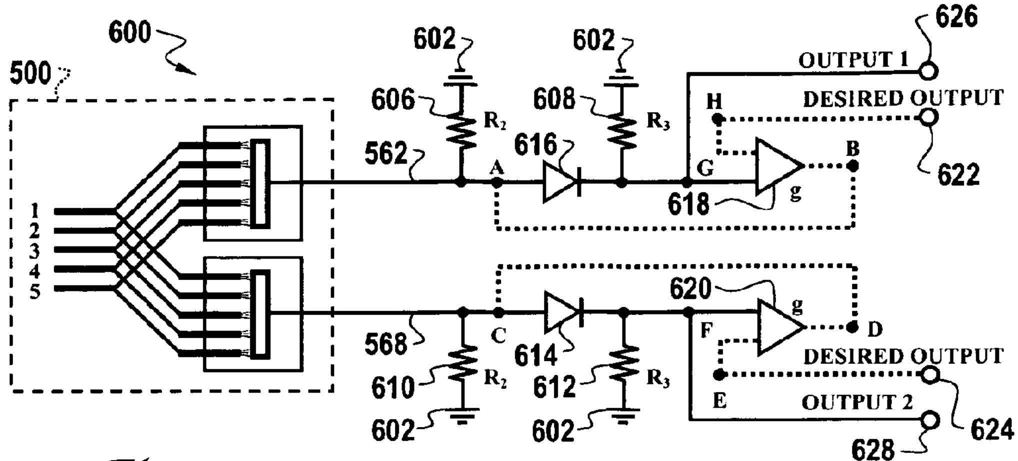
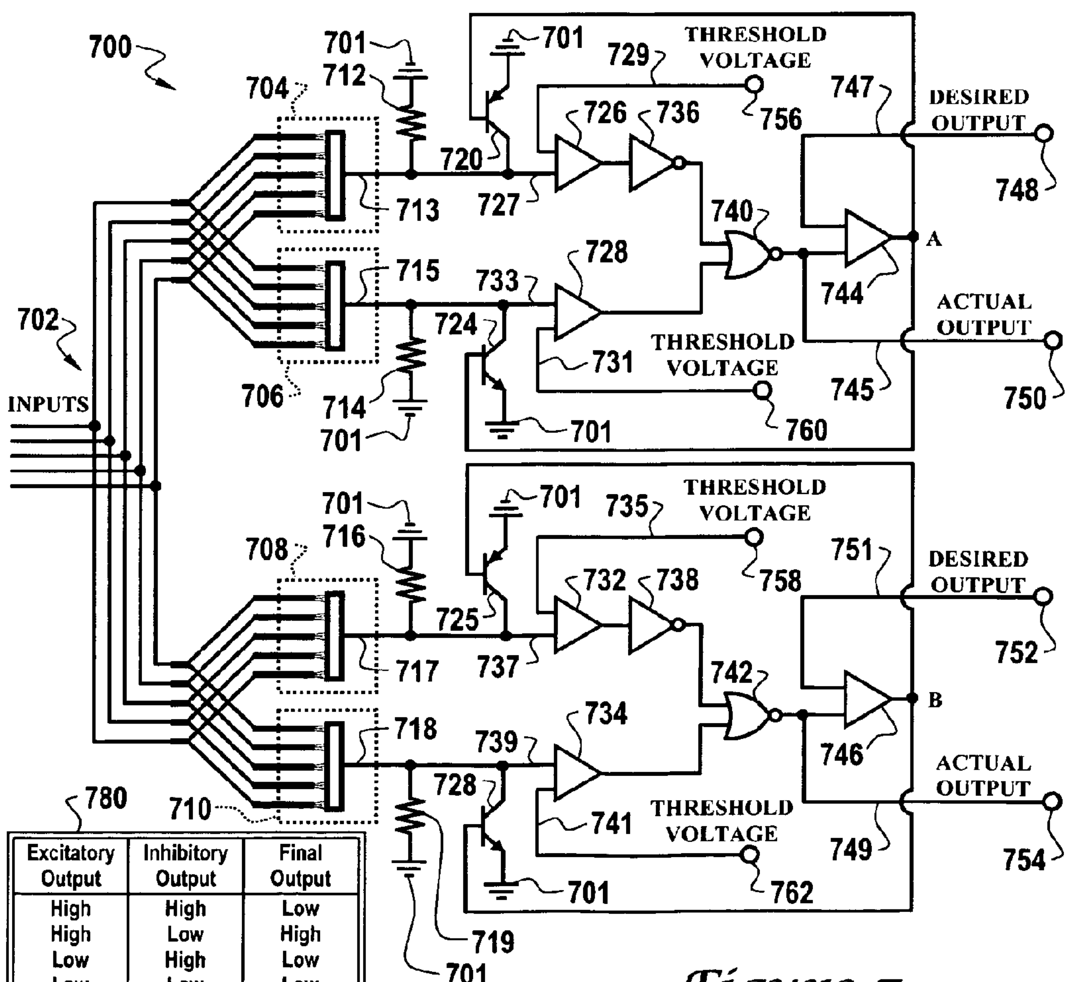


Figure 6



Excitatory Output	Inhibitory Output	Final Output
High	High	Low
High	Low	High
Low	High	Low
Low	Low	Low

Figure 7

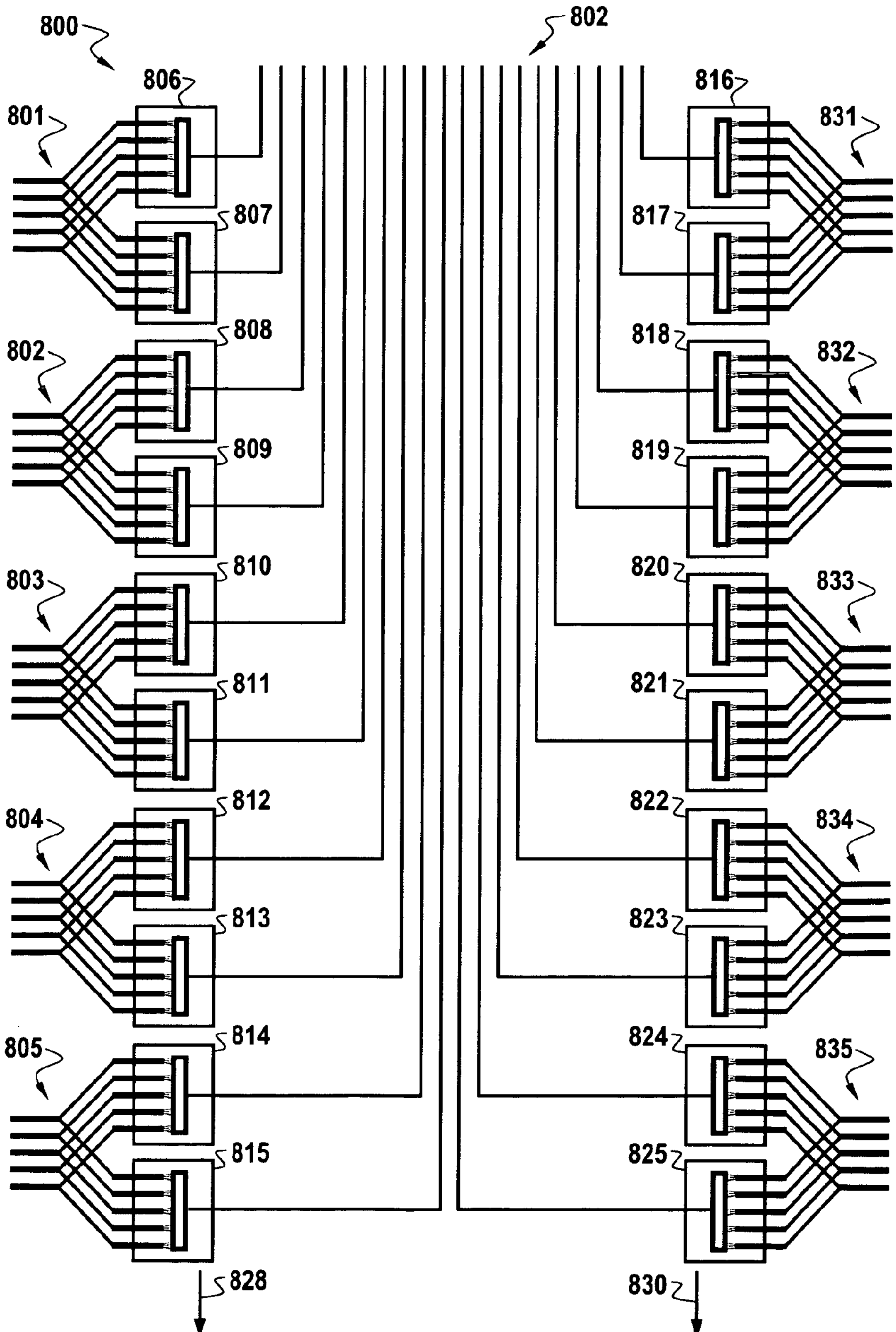

800
802
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
828
830
831
832
833
834
835

Figure 8

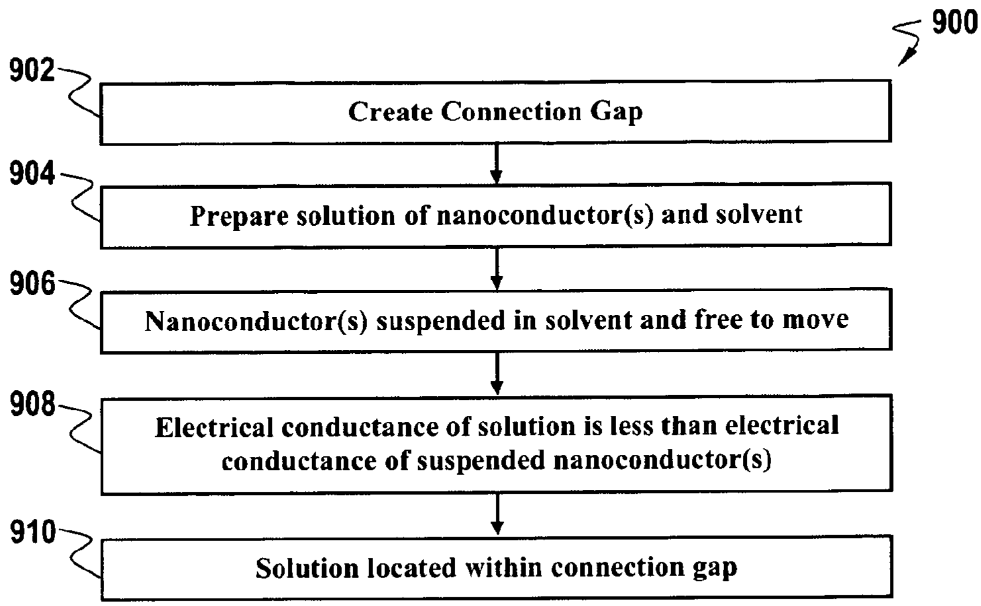


Figure 9

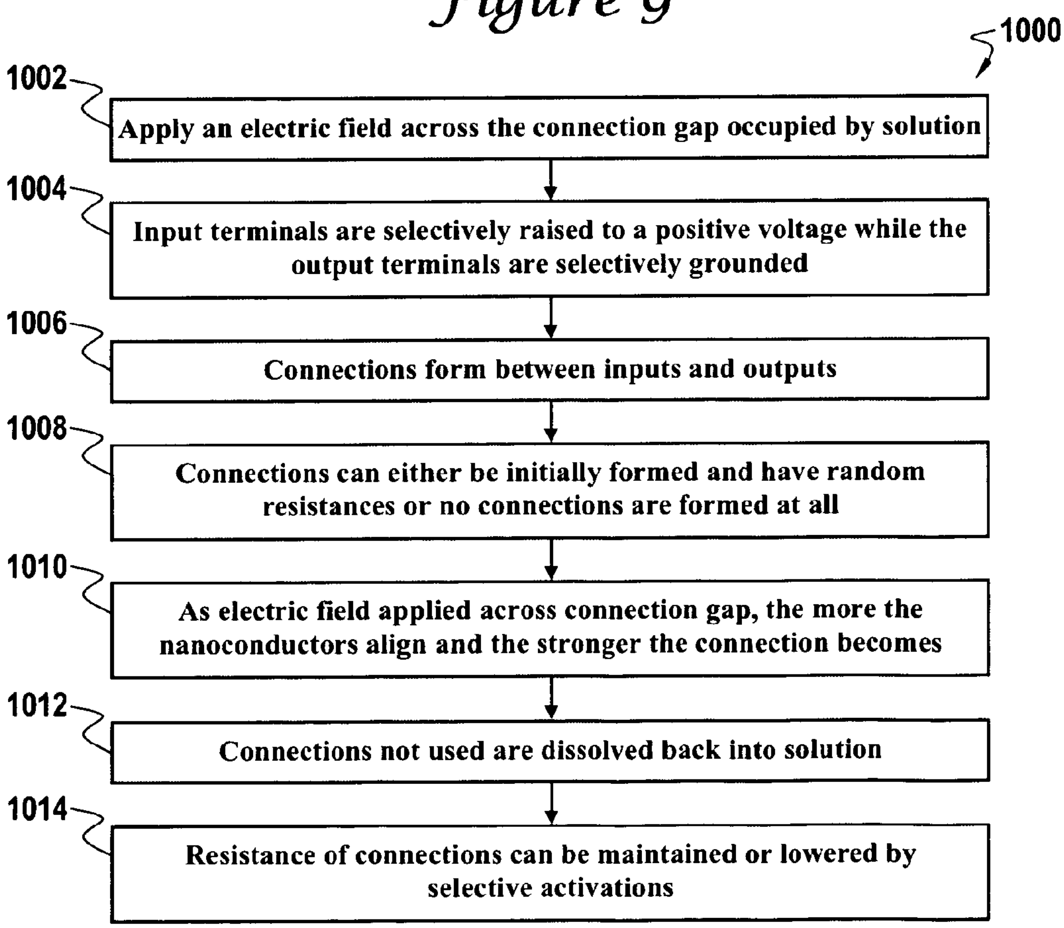


Figure 10

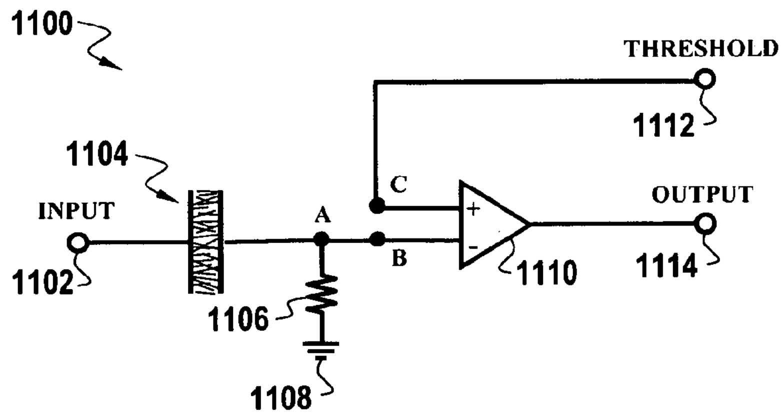


Figure 11

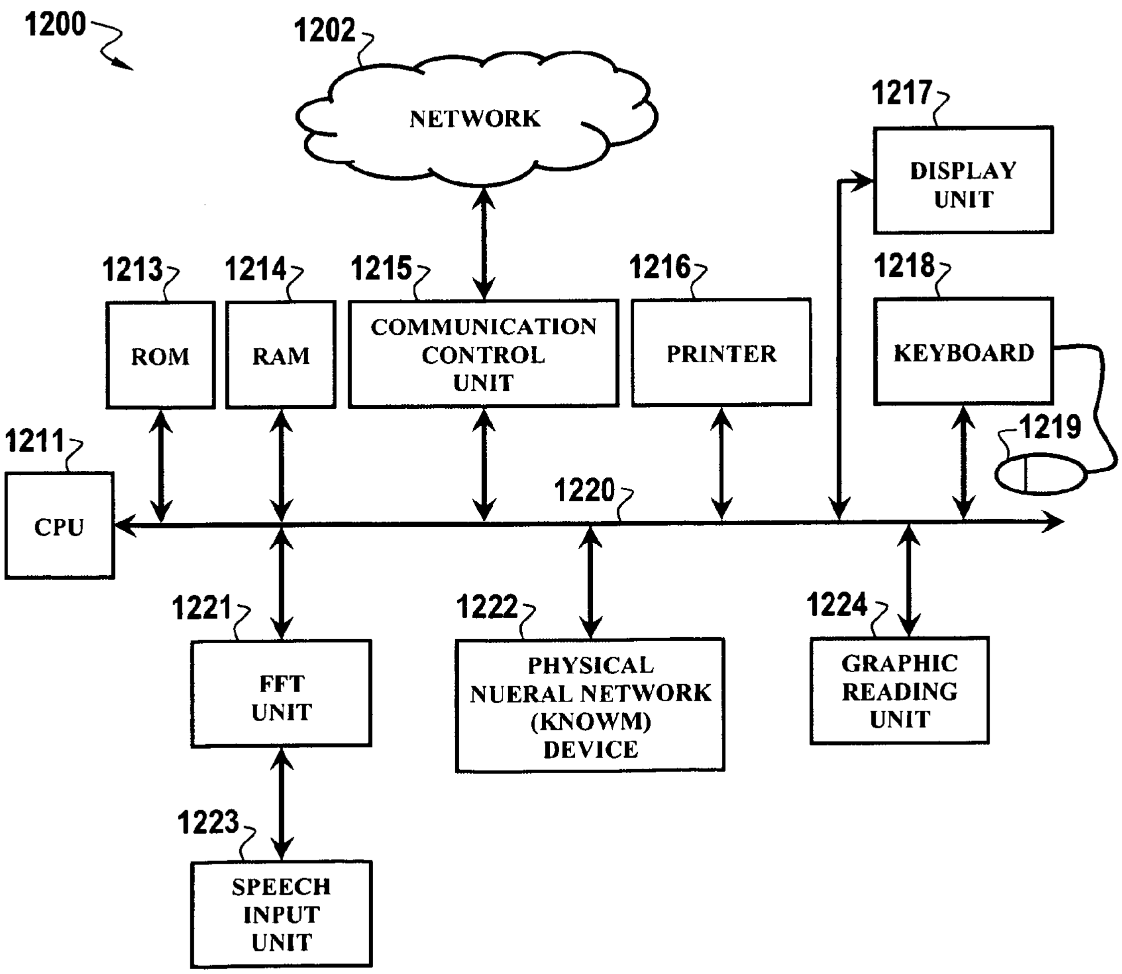


Figure 12

TEMPORAL SUMMATION DEVICE UTILIZING NANOTECHNOLOGY

CROSS REFERENCE TO RELATED PATENT APPLICATION

This patent application is a continuation of U.S. patent application Ser. No. 10/095,273 entitled "Physical Neural Network Design Incorporating Nanotechnology," which was filed on Mar. 12, 2002, the disclosure of which is incorporated herein by reference in its entirety. U.S. patent application Ser. No. 10/095,272 issued as U.S. Pat. No. 6,889,216 on May 3, 2005.

TECHNICAL FIELD

Embodiments generally relate to nanotechnology. Embodiments also relate to neural networks and neural computing systems and methods thereof. Embodiments also relate to temporal summation devices.

BACKGROUND

Neural networks are computational systems that permit computers to essentially function in a manner analogous to that of the human brain. Neural networks do not utilize the traditional digital model of manipulating **0**'s and **1**'s. Instead, neural networks create connections between processing elements, which are equivalent to neurons of a human brain. Neural networks are thus based on various electronic circuits that are modeled on human nerve cells (i.e., neurons). Generally, a neural network is an information-processing network, which is inspired by the manner in which a human brain performs a particular task or function of interest. Computational or artificial neural networks are thus inspired by biological neural systems. The elementary building block of biological neural systems is of course the neuron, the modifiable connections between the neurons, and the topology of the network.

In general, artificial neural networks are systems composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. The computational elements, or nodes, are connected via variable weights that are typically adapted during use to improve performance. Thus, in solving a problem, neural net models can explore many competing hypothesis simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights.

In a neural network, "neuron-like" nodes can output a signal based on the sum of their inputs, the output being the result of an activation function. In a neural network, there exists a plurality of connections, which are electrically coupled among a plurality of neurons. The connections serve as communication bridges among of a plurality of neurons coupled thereto. A network of such neuron-like nodes has the ability to process information in a variety of useful ways. By adjusting the connection values between neurons in a network, one can match certain inputs with desired outputs.

A number of software simulations of neural networks have been developed. Because software simulations are performed on conventional sequential computers, however, they do not take advantage of the inherent parallelism of neural network architectures. Consequently, they are relatively slow. One frequently used measurement of the speed of a neural network processor is the number of interconnections it can perform per second. For example, the fastest software simulations available can perform up to about 18 million interconnects per second. Such speeds, however, currently require expensive super computers to achieve. Even so, 18 million interconnects per second is still too slow to perform many pattern classification tasks in real time.

The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions. Due to the difficulties in building such highly interconnected processors, the currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and built to implement one particular neural network architecture and are not easily, if at all, reconfigurable to implement different architectures. A true physical neural network (i.e., artificial neural network) chip, for example, has not yet been designed and successfully implemented.

The term "Nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

Based on the foregoing, it is believed that a physical neural network which incorporates nanotechnology is a solution to the problems encountered by prior art neural network solutions. Additionally, it is believed that a temporal summation device can be constructed based on nanotechnology and utilized either as an individual component for temporal summation purposes, or in association with physical neural networks, including artificial neurons and components thereof as described herein.

BRIEF SUMMARY

The following summary is provided to facilitate an understanding of some of the innovative features unique to the embodiments, and is not intended to be a full description. A full appreciation of the various aspects of the embodiments can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

It is, therefore, one aspect of the present invention to provide for a temporal summation device that is formed based on nanotechnology.

It is another aspect of the present invention to provide a physical neural network, which can be formed from a plurality of interconnected nanoconnections or nanoconnectors.

It is a further aspect of the present invention to provide neuron like nodes, which can be formed and implemented utilizing nanotechnology;

It is also an aspect of the present invention to provide a physical neural network that can be formed from one or more neuron-like nodes.

It is yet a further aspect of the present invention to provide a physical neural network, which can be formed from a plurality of nanoconductors, such as, for example, nanowires and/or nanotubes.

The above and other aspects can be achieved as is now described. A physical neural network based on nanotechnology is disclosed herein, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, connected to a plurality of interconnected nanoconnections. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes can be strengthened or weakened according to an application of an electric field. Alignment has also been observed with a magnetic field, but electric fields are generally more practical. Note that the connection network is generally associated with one or more of the neuron-like nodes.

The output signal is generally based on a threshold below which the output signal is not generated and above which the output signal is generated. The transition from zero output to high output need not necessarily be abrupt or non linear. The connection network comprises a number of layers of nanoconnections, wherein the number of layers is generally equal to a number of desired outputs from the connection network. The nanoconnections are formed without influence from disturbances resulting from other nanoconnections thereof. Such nanoconnections may be formed from an electrically conducting material. The electrically conducting material can be selected such that a dipole is induced in the electrically conducting material in the presence of an electric field. Such a nanoconnection may comprise a nanoconductor.

The connection network itself may comprise a connection network structure having a connection gap formed therein, and a solution located within the connection gap, such that the solution comprises a solvent or suspension and one or more nanoconductors. Preferably, a plurality of nanoconductors is present in the solution (i.e., mixture). Note that such a solution may comprise a liquid and/or gas. An electric field can then be applied across the connection gap to permit the alignment of one or more of the nanoconductors within the connection gap. The nanoconductors can be suspended in the solvent, or can lie at the bottom of the connection gap on the surface of the chip. Studies have shown that nanotubes can align both in the suspension and/or on the surface of the gap. The electrical conductance of the mixture is less than the electrical conductance of the nanoconductors within the solution.

The nanoconductors within the connection gap thus experience an increased alignment in accordance with an increase in the electric field applied across the connection gap. Thus, nanoconnections of the neuron-like node that are utilized most frequently by the neuron-like node become stronger with each use thereof. The nanoconnections that are utilized least frequently become increasingly weak and eventually dissolve back into the solution. The nanoconnections may or may not comprise a resistance, which can be raised or lowered by a selective activation of a nanoconnection. They can be configured as nanoconductors such as, for example, a nanotube or nanowire. An example of a nanotube, which may be implemented in accordance with the invention described herein, is a carbon nanotube, nanowire and/or other nanoparticle. Additionally, such nanoconnections may be configured as a negative connection associated with the neuron-like node.

The components described herein can be arranged to provide for a temporal summation device that is composed of one or more nanoconnections having an input and an output thereof, wherein an input signal provided to the input causes one or more of the nanoconnection to experience an increase in connection strength thereof over time. Additionally, a voltage divider is formed by the nanoconnection(s) and a resistor connected to the output of the nanoconnection(s), such that the voltage divider provides a voltage at the output the nanoonnection(s) that is in direct proportion to the connection strength of nanoconnection(s). An amplifier is also connected to the voltage divider, wherein when the voltage provided by the voltage divider attains a desired threshold voltage, the amplifier attains a high voltage output thereby providing a temporal summation device thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a graph illustrating a typical activation function that can be implemented in accordance with one embodiment;

FIG. **2** illustrates a schematic diagram illustrating a diode configuration as a neuron, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanowires between two electrodes, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections formed between two electrodes, in accordance with a preferred embodiment;

FIG. **4** illustrates a block diagram illustrating a plurality of connections between inputs and outputs of a physical neural network, in accordance with a preferred embodiment;

FIG. **5** illustrates a schematic diagram of a physical neural network that can be created without disturbances, in accordance with a preferred embodiment;

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **8** illustrates a schematic diagram of a chip layout for a connection network that may be implemented in accordance with an alternative embodiment;

FIG. **9** illustrates a flow chart of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment;

FIG. **10** illustrates a flow chart of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment;

FIG. **11** illustrates a schematic diagram of a circuit illustrating temporal summation within a neuron, in accordance with a preferred embodiment; and

FIG. **12** illustrates a block diagram illustrating a pattern recognition system, which may be implemented with a physical neural network device, in accordance with a preferred embodiment.

DETAILED DESCRIPTION

The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate one or more embodiments.

The physical neural network described and disclosed herein is different from prior art forms of neural networks in that the disclosed physical neural network does not require a computer simulation for training, nor is its architecture based on any current neural network hardware device. The design of the physical neural network described herein with respect to particular embodiments is actually quite "organic". Such a physical neural network is generally fast and adaptable, no matter how large such a physical neural network becomes. The physical neural network described herein can be referred to generically as a Knowm. The terms "physical neural network" and "Knowm" can be utilized interchangeably to refer to the same device, network, or structure.

Network orders of magnitude larger than current VSLI neural networks can be built and trained with a standard computer. One consideration for a Knowm is that it must be large enough for its inherent parallelism to shine through. Because the connection strengths of such a physical neural network are dependent on the physical movement of nano-connections thereof, the rate at which a small network can learn is generally very small and a comparable network simulation on a standard computer can be very fast. On the other hand, as the size of the network increases, the time to train the device does not change. Thus, even if the network takes a full second to change a connection value a small amount, if it does the same to a billion connections simultaneously, then its parallel nature begins to express itself.

A physical neural network (i.e., a Knowm) must have two components to function properly. First, the physical neural network must have one or more neuron-like nodes that sum a signal and output a signal based on the amount of input signal received. Such a neuron-like node is generally non-linear in its output. In other words, there should be a certain threshold for input signals, below which nothing is output and above which a constant or nearly constant output is generated or allowed to pass. This is a very basic requirement of standard software-based neural networks, and can be accomplished by an activation function. The second requirement of a physical neural network is the inclusion of a connection network composed of a plurality of interconnected connections (i.e., nanoconnections). Such a connection network is described in greater detail herein.

FIG. **1** illustrates a graph **100** illustrating a typical activation function that can be implemented in accordance with one embodiment. Note that the activation function need not be non-linear, although non-linearity is generally desired for learning complicated input-output relationships. The activation function depicted in FIG. **1** comprises a linear function, and is shown as such for general edification and illustrative purposes only. As explained previously, an activation function may also be non-linear.

As illustrated in FIG. **1**, graph **100** includes a horizontal axis **104** representing a sum of inputs, and a vertical axis **102** representing output values. A graphical line **106** indicates threshold values along a range of inputs from approximately −10 to +10 and a range of output values from approximately 0 to 1. As more neural networks (i.e., active inputs) are established, the overall output as indicated at line **105** climbs until the saturation level indicated by line **106** is attained. If a connection is not utilized, then the level of output (i.e., connection strength) begins to fade until it is revived. This phenomenon is analogous to short term memory loss of a human brain. Note that graph **100** is presented for generally illustrative and edification purposes only and is not considered a limiting feature of the embodiments.

In a Knowm network, the neuron-like node can be configured as a standard diode-based circuit, the diode being the most basic semiconductor electrical component, and the signal it sums may be a voltage. An example of such an arrangement of circuitry is illustrated in FIG. **2**, which generally illustrates a schematic diagram illustrating a diode-based configuration as a neuron **200**, in accordance with a preferred embodiment. Those skilled in the art can appreciate that the use of such a diode-based configuration is not considered a limitation of the embodiments, but merely represents one potential arrangement in which the embodiments may be implemented.

Although a diode may not necessarily be utilized, its current versus voltage characteristics are non-linear when used with associated resistors and similar to the relationship depicted in FIG. **1**. The use of a diode as a neuron is thus not a limiting feature, but is only referenced herein with respect to a preferred embodiment. The use of a diode and associated resistors with respect to a preferred embodiment simply represents one potential "neuron" implementation. Such a configuration can be said to comprise an artificial neuron. It is anticipated that other devices and components may be utilized instead of a diode to construct a physical neural network and a neuron-like node (i.e., artificial neuron), as indicated here.

Thus, neuron **200** comprises a neuron-like node that may include a diode **206**, which is labeled D_1, and a resistor **204**, which is labeled R_2. Resistor **204** is connected to a ground **210** and an input **205** of diode **206**. Additionally, a resistor **202**, which is represented as a block and labeled R_1 can be connected to input **205** of diode **206**. Block **202** includes an input **212**, which comprises an input to neuron **200**. A resistor **208**, which is labeled R_3, is also connected to an output **214** of diode **206**. Additionally, resistor **208** is coupled to ground **210**. Diode **206** in a physical neural network is analogous to a neuron of a human brain, while an associated connection formed thereof, as explained in greater detail herein, is analogous to a synapse of a human brain.

As depicted in FIG. **2**, the output **214** is determined by the connection strength of R_1 (i.e., resistor **202**). If the strength of R_1's connection increases (i.e., the resistance decreases), then the output voltage at output **214** also increases. Because diode **206** conducts essentially no current until its threshold voltage (e.g., approximately 0.6V for silicon) is attained, the output voltage will remain at zero until R_1 conducts enough current to raise the pre-diode voltage to approximately 0.6V. After 0.6V has been achieved, the output voltage at output **214** will increase linearly. Simply adding extra diodes in series or utilizing different diode types may increase the threshold voltage.

An amplifier may also be added to the output **214** of diode **206** so that the output voltage immediately saturates at the diode threshold voltage, thus resembling a step function, until a threshold value and a constant value above the threshold is attained. R_3 (i.e., resistor **208**) functions generally as a bias for diode **206** (i.e., D_1) and should generally be about 10 times larger than resistor **204** (i.e., R_2). In the circuit configuration illustrated in FIG. **2**, R_1 can actually be configured as a network of connections composed of many inter-connected conducting nanowires (i.e., see FIG. **3**). As explained previously, such connections are analogous to the synapses of a human brain.

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections **304** formed between two electrodes, in accordance with a preferred embodiment. Nanoconnections **304** (e.g., nanoconductors) depicted in FIG. **3** are generally located between input **302** and output **306**. The network of nanoconnections depicted in FIG. **3** can be implemented as a network of nanoconductors. Examples of nanoconductors include devices such as, for example, nanowires, nanotubes, and nanoparticles.

Nanoconnections **304**, which are analogous to the synapses of a human brain, are preferably composed of electrical conducting material (i.e., nanoconductors). It should be appreciated by those skilled in the art that such nanoconductors can be provided in a variety of shapes and sizes without departing from the teachings herein. For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles.

For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles. For example, each nanoconnection within a connection network may be formed from as a chain of carbon nanoparticles. In "Self-assembled chains of graphitized carbon nanoparticles" by Bezryadin et al., Applied Physics Letters, Vol. 74, No. 18, pp. 2699–2701, May 3, 1999, for example, a technique is reported, which permits the self-assembly of conducting nanoparticles into long continuous chains. Thus, nanoconductors which are utilized to form a physical neural network (i.e., Knowm) could be formed from such nanoparticles. It can be appreciated that the Bezryadin et al is referred to herein for general edification and illustrative purposes only and is not considered to limit the embodiments.

It can be appreciated that a connection network as disclosed herein may be composed from a variety of different types of nanoconductors. For example, a connection network may be formed from a plurality of nanoconductors, including nanowires, nanotubes and/or nanoparticles. Note that such nanowires, nanotubes and/or nanoparticles, along with other types of nanoconductors can be formed from materials such as carbon or silicon. For example, carbon nanotubes may comprise a type of nanotube that can be utilized in accordance with one or more embodiments.

As illustrated in FIG. **3**, nanoconnections **304** comprise a plurality of interconnected nanoconnections, which from this point forward, can be referred to generally as a "connection network." An individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire, a nanotube, nanoparticles(s), or any other nanoconducting structures. Nanoconnections **304** may comprise a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections **304** may be formed from a plurality of interconnected nanoparticles. A connection network is thus not one connection between two electrodes, but a plurality of connections between inputs and outputs. Nanotubes, nanowires, nanoparticles and/or other nanoconducting structures may be utilized, of course, to construct nanoconnections **304** between input **302** and input **306**. Although a single input **302** and a single input **306** is depicted in FIG. **3**, it can be appreciated that a plurality of inputs and a plurality of outputs may be implemented in accordance with the embodiments, rather than simply a single input **302** or a single output **306**.

FIG. **4** illustrates a block diagram illustrating a plurality of nanoconnections **414** between inputs **404**, **406**, **408**, **410**, **412** and outputs **416** and **418** of a physical neural network, in accordance with a preferred embodiment. Inputs **404**, **406**, **408**, **410**, and **412** can provide input signals to connections **414**. Output signals can then be generated from connections **414** via outputs **416** and **418**. A connection network can therefore be configured from the plurality of connections **414**. Such a connection network is generally associated with one or more neuron-like nodes.

The connection network also comprises a plurality of interconnected nanoconnections, wherein each nanoconnection thereof is strengthened or weakened according to an application of an electric field. A connection network is not possible if built in one layer because the presence of one connection can alter the electric field so that other connections between adjacent electrodes could not be formed. Instead, such a connection network can be built in layers, so that each connection thereof can be formed without being influenced by field disturbances resulting from other connections. This can be seen in FIG. **5**.

FIG. **5** illustrates a schematic diagram of a physical neural network **500** that can be created without disturbances, in accordance with a preferred embodiment. Physical neural network **500** is composed of a first layer **558** and a second layer **560**. A plurality of inputs **502**, **504**, **506**, **508**, and **510** are respectively provided to layers **558** and **560** respectively via a plurality of input lines **512**, **514**, **516**, **518**, and **520** and a plurality of input lines **522**, **524**, **526**, **528**, and **530**. Input lines **512**, **514**, **516**, **518**, and **520** are further coupled to input lines **532**, **534**, **536**, **538**, and **540** such that each line **532**, **534**, **536**, **538**, and **540** is respectively coupled to nanoconnections **572**, **574**, **576**, **578**, and **580**. Thus, input line **532** is connected to nanconnections **572**. Input line **534** is connected to nanoconnections **574**, and input line **536** is connected to nanoconnections **576**. Similarly, input line **538** is connected to nanconnections **578**, and input line **540** is connected to nanoconnections **580**.

Nanconnections **572**, **574**, **576**, **578**, and **580** may comprise nanoconductors such as, for example, nanotubes and/or nanowires. Nanoconnections **572**, **574**, **576**, **578**, and **580** thus comprise one or more nanoconductors. Additionally, input lines **522**, **524**, **526**, **528**, and **530** are respectively coupled to a plurality of input lines **542**, **544**, **546**, **548** and **550**, which are in turn each respectively coupled to nanoconnections **582**, **584**, **586**, **588**, and **590**. Thus, for example, input line **542** is connected to nanoconnections **582**, while input line **544** is connected to nanoconnections **584**. Similarly, input line **546** is connected to nanoconnections **586** and input line **548** is connected to nanoconnections **588**. Additionally, input line **550** is connected to nanconnections **590**. Box **556** and **554** generally represent simply the output and are thus illustrated connected to outputs **562** and **568**. In other words, outputs **556** and **554** respectively comprise outputs **562** and **568**. The aforementioned input lines and associated components thereof actually comprise physical electronic components, including conducting input and output lines and physical nanoconnections, such as nanotubes and/or nanowires.

Thus, the number of layers **558** and **560** equals the number of desired outputs **562** and **568** from physical neural network **500**. In the previous two figures, every input was potentially connected to every output, but many other configurations are possible. The connection network can be made of any electrically conducting material, although the physics of it requires that they be very small so that they will align with a practical voltage. Carbon nanotubes or any conductive nanowire can be implemented in accordance with the physical neural network described herein. Such

components can form connections between electrodes by the presence of an electric field. For example, the orientation and purification of carbon nanotubes has been demonstrated using ac electrophoresis in isopropyl alcohol, as indicated in "Orientation and purification of carbon nanotubes using ac electrophoresis" by Yamamoto et al., J. Phys. D: Applied Physics, 31 (1998), 34–36. Additionally, an electric-field assisted assembly technique used to position individual nanowires suspended in an electric medium between two electrodes defined lithographically on an SiO_2 substrate is indicated in "Electric-field assisted assembly and alignment of metallic nanowires," by Smith et al., Applied Physics Letters, Vol. 77, Num. 9, Aug. 28, 2000. Such references are referred to herein for edification and illustrative purposes only.

The only general requirements for the conducting material utilized to configure the nanoconductors are that such conducting material should preferably conduct electricity, and a dipole should preferably be induced in the material when in the presence of an electric field. Alternatively, the nanoconductors utilized in association with the physical neural network described herein can be configured to include a permanent dipole that is produced by a chemical means, rather than a dipole that is induced by an electric field.

Therefore, it should be appreciated by those skilled in the art that a connection network could also be comprised of other conductive particles that may be developed or found useful in the nanotechnology arts. For example, carbon particles (or "dust") may also be used as nanoconductors in place of nanowires or nanotubes. Such particles may include bearings or granule-like particles.

A connection network can be constructed as follows: A voltage is applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture could be made of many things. The only requirements are that the conducting wires must be suspended in the solvent, either dissolved or in some sort of suspension, free to move around; the electrical conductance of the substance must be less than the electrical conductance of the suspended conducting wire; and the viscosity of the substance should not be too much so that the conducting wire cannot move when an electric field is applied.

The goal for such a connection network is to develop a network of connections of just the right values so as to satisfy the particular signal-processing requirement—exactly what a neural network does. Such a connection network can be constructed by applying a voltage across a space occupied by the mixture mentioned. To create the connection network, the input terminals are selectively raised to a positive voltage while the output terminals are selectively grounded. Thus, connections can gradually form between the inputs and outputs. The important requirement that makes the physical neural network functional as a neural network is that the longer this electric field is applied across a connection gap, or the greater the frequency or amplitude, the more nanotubes and/or nanowires and/or particles align and the stronger the connection thereof becomes. Thus, the connections that are utilized most frequently by the physical neural network become the strongest.

The connections can either be initially formed and have random resistances or no connections may be formed at all. By initially forming random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up from scratch. Depending on the rate of connection decay, having initial random connections could prove faster, although not necessarily. The connection network can adapt itself to the requirements of a given situation regardless of the initial state of the connections. Either initial condition will work, as connections that are not used will "dissolve" back into solution. The resistance of the connection can be maintained or lowered by selective activations of the connection. In other words, if the connection is not used, it will fade away, analogous to the connections between neurons in a human brain. The temperature of the solution can also be maintained at a particular value so that the rate that connections fade away can be controlled. Additionally an electric field can be applied perpendicular to the connections to weaken them, or even erase them out altogether (i.e., as in clear, zero, or reformatting of a "disk").

The nanoconnections may or may not be arranged in an orderly array pattern. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, due to the non-linear nature of neural networks, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network **600** that can be implemented in accordance with an alternative embodiment. Note that in FIGS. **5** and **6**, like parts are indicated by like reference numerals. Thus, physical neural network **600** can be configured, based on physical neural network **500** illustrated in FIG. **5**. In FIG. **6**, inputs **1**, **2**, **3**, **4**, and **5** are indicated, which are respectively analogous to inputs **502**, **504**, **506**, **508**, and **510** illustrated in FIG. **5**. Outputs **562** and **568** are provided to a plurality of electrical components to create a first output **626** (i.e., Output **1**) and a second output **628** (i.e., Output **2**). Output **562** is tied to a resistor **606**, which is labeled R2 and a diode **616** at node A. Output **568** is tied to a resistor **610**, which is also labeled R2 and a diode **614** at node C. Resistors **606** and **610** are each tied to a ground **602**.

Diode **616** is further coupled to a resistor **608**, which is labeled R3, and first output **626**. Additionally, resistor **608** is coupled to ground **602** and an input to an amplifier **618**. An output from amplifier **618**, as indicated at node B and dashed lines thereof, can be tied back to node A. A desired output **622** from amplifier **618** is coupled to amplifier **618** at node H. Diode **614** is coupled to a resistor **612** at node F. Note that resistor **612** is labeled R3. Node F is in turn coupled to an input of amplifier **620** and to second output **628** (i.e., Output **2**). Diode **614** is also connected to second output **628** and an input to amplifier **620** at second output **628**. Note that second output **628** is connected to the input to amplifier **620** at node F. An output from amplifier **620** is further coupled to node D, which in turn is connected to node C. A desired output **624**, which is indicated by a dashed line in FIG. **6**, is also coupled to an input of amplifier **620** at node E.

In FIG. **6**, the training of physical neural network **600** can be accomplished utilizing, for example, op-amp devices (e.g., amplifiers **618** and **620**). By comparing an output (e.g., first output **626**) of physical neural network **600** with a desired output (e.g., desired output **622**), the amplifier (e.g., amplifier **618**) can provide feedback and selectively

strengthen connections thereof. For instance, suppose it is desired to output a voltage of +V at first output **626** (i.e., Output **1**) when inputs **1** and **4** are high. When inputs **1** and **4** are taken high, also assume that first output **626** is zero. Amplifier **618** can then compare the desired output (+V) with the actual output (**0**) and output −V. In this case, −V is equivalent to ground.

The op-amp outputs and grounds the pre-diode junction (i.e., see node A) and causes a greater electric field across inputs **1** and **4** and the layer **1** output. This increased electric field (larger voltage drop) can cause the nanoconductors in the solution between the electrode junctions to align themselves, aggregate, and form a stronger connection between the 1 and 4 electrodes. Feedback can continue to be applied until output of physical neural network **600** matches the desired output. The same procedure can be applied to every output.

In accordance with the aforementioned example, assume that Output **1** was higher than the desired output (i.e., desired output **622**). If this were the case, the op-amp output can be +V and the connection between inputs **1** and **4** and layer one output can be raised to +V. Columbic repulsions between the nanoconductors can force the connection apart, thereby weakening the connection. The feedback will then continue until the desired output is obtained. This is just one training mechanism. One can see that the training mechanism does not require any computations, because it is a simple feedback mechanism.

Such a training mechanism, however, may be implemented in many different forms. Basically, the connections in a connection network must be able to change in accordance with the feedback provided. In other words, the very general notion of connections being strengthened or connections being weakened in a physical system is the essence of a physical neural network (i.e., Knowm). Thus, it can be appreciated that the training of such a physical neural network may not require a "CPU" to calculate connection values thereof. The Knowm can adapt itself. Complicated neural network solutions could be implemented very rapidly "on the fly", much like a human brain adapts as it performs.

The physical neural network disclosed herein thus has a number of broad applications. The core concept of a Knowm, however, is basic. The very basic idea that the connection values between electrode junctions by nanoconductors can be used in a neural network devise is all that required to develop an enormous number of possible configurations and applications thereof.

Another important feature of a physical neural network is the ability to form negative connections. This is an important feature that makes possible inhibitory effects useful in data processing. The basic idea is that the presence of one input can inhibit the effect of another input. In artificial neural networks as they currently exist, this is accomplished by multiplying the input by a negative connection value. Unfortunately, with a Knowm-based device, the connection may only take on zero or positive values under such a scenario.

In other words, either there can be a connection or no connection. A connection can simulate a negative connection by dedicating a particular connection to be negative, but one connection cannot begin positive and through a learning process change to a negative connection. In general, if starts positive, it can only go to zero. In essence, it is the idea of possessing a negative connection initially that results in the simulation, because this does not occur in a human brain. Only one type of signal travels through axon/dendrites in a human brain. That signal is transferred into the flow of a neurotransmitter whose effect on the postsynaptic neuron can be either excitatory or inhibitory, depending on the neuron.

One method for solving this problem is to utilize two sets of connections for the same output, having one set represent the positive connections and the other set represent the negative connections. The output of these two layers can be compared, and the layer with the greater output will output either a high signal or a low signal, depending on the type of connection set (inhibitory or excitatory). This can be seen in FIG. **7**.

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network **700** that can be implemented in accordance with an alternative embodiment. Physical neural network **700** thus comprises a plurality of inputs **702** (not necessarily binary) which are respectively fed to layers **704**, **706**, **708**, and **710**. Each layer is analogous to the layers depicted earlier, such as for example layers **558** and **560** of FIG. **5**. An output **713** of layer **704** can be connected to a resistor **712**, a transistor **720** and a first input **727** of amplifier **726**. Transistor **720** is generally coupled between ground **701** and first input **727** of amplifier **726**. Resistor **712** is connected to a ground **701**. Note that ground **701** is analogous to ground **602** illustrated in FIG. **6** and ground **210** depicted in FIG. **2**. A second input **729** of amplifier **726** can be connected to a threshold voltage **756**. The output of amplifier **726** can in turn be fed to an inverting amplifier **736**.

The output of inverting amplifier **736** can then be input to a NOR device **740**. Similarly, an output **716** of layer **706** may be connected to resistor **714**, transistor **733** and a first input **733** of an amplifier **728**. A threshold voltage **760** is connected to a second input **737** of amplifier **728**. Resistor **714** is generally coupled between ground **701** and first input **733** of amplifier **728**. Note that first input **733** of amplifier **728** is also generally connected to an output **715** of layer **706**. The output of amplifier **728** can in turn be provided to NOR device **740**. The output from NOR device **740** is generally connected to a first input **745** of an amplifier **744**. An actual output **750** can be taken from first input **745** to amplifier **744**. A desired output **748** can be taken from a second input **747** to amplifier **744**. The output from amplifier **744** is generally provided at node A, which in turn is connected to the input to transistor **720** and the input to transistor **724**. Note that transistor **724** is generally coupled between ground **701** and first input **733** of amplifier **728**. The second input **731** of amplifier **728** can produce a threshold voltage **760**.

Layer **708** provides an output **717** that can be connected to resistor **716**, transistor **725** and a first input **737** to an amplifier **732**. Resistor **716** is generally coupled between ground **701** and the output **717** of layer **708**. The first input **737** of amplifier **732** is also electrically connected to the output **717** of layer **708**. A second input **735** to amplifier **732** may be tied to a threshold voltage **758**. The output from amplifier **732** can in turn be fed to an inverting amplifier **738**. The output from inverting amplifier **738** may in turn be provided to a NOR device **742**. Similarly, an output **718** from layer **710** can be connected to a resistor **719**, a transistor **728** and a first input **739** of an amplifier **734**. Note that resistor **719** is generally coupled between node **701** and the output **719** of layer **710**. A second input **741** of amplifier **734** may be coupled to a threshold voltage **762**. The output from of NOR device **742** is generally connected to a first input **749** of an amplifier **746**. A desired output **752** can be taken from a second input **751** of amplifier **746**. An actual output **754** can be taken from first input **749** of amplifier **746**.

The output of amplifier **746** may be provided at node B, which in turn can be tied back to the respective inputs to transistors **725** and **728**. Note that transistor **725** is generally coupled between ground **701** and the first input **737** of amplifier **732**. Similarly, transistor **728** is generally connected between ground **701** and the first input **739** of amplifier **734**.

Note that transistors **720**, **724**, **725** and/or **728** each can essentially function as a switch to ground. A transistor such as, for example, transistor **720**, **724**, **725** and/or **728** may comprise a field-effect transistor (FET) or another type of transistor, such as, for example, a single-electron transistor (SET). Single-electron transistor (SET) circuits are essential for hybrid circuits combining quantum SET devices with conventional electronic devices. Thus, SET devices and circuits may be adapted for use with the physical neural network of the embodiments. This is particularly important because as circuit design rules begin to move into regions of the sub-100 nanometer scale, where circuit paths are only 0.001 of the thickness of a human hair, prior art device technologies will begin to fail, and current leakage in traditional transistors will become a problem. SET offers a solution at the quantum level, through the precise control of a small number of individual electrons. Transistors such as transistors **720**, **724**, **725** and/or **728** can also be implemented as carbon nanotube transistors.

A truth table for the output of circuit **700** is illustrated at block **780** in FIG. **7**. As indicated at block **780**, when an excitatory output is high and the inhibitory output is also high, the final output is low. When the excitatory output is high and the inhibitory output is low, the final output is high. Similarly, when the excitatory output is low and the inhibitory output is high, the final output is low. When the excitatory output is low and the inhibitory output is also low, the final output is low. Note that layers **704** and **708** may thus comprise excitatory connections, while layers **706** and **710** may comprise inhibitory connections.

For every desired output, two sets of connections are used. The output of a two-diode neuron can be fed into an op-amp (e.g., a comparator). If the output that the op-amp receives is low when it should be high, the op-amp outputs a low signal. This low signal can cause the transistors (e.g., transistors **720**, **725**) to saturate and ground out the pre-diode junction for the excitatory diode. Such a scenario can cause, as indicated previously, an increase in the voltage drop across those connections that need to increase their strength. Note that only those connections going to the excitatory diode are strengthened. Likewise, if the desired output were low when the actual output was high, the op-amp can output a high signal. This can cause the inhibitory transistor (e.g., an NPN transistor) to saturate and ground out the neuron junction of the inhibitory connections. Those connections going to the inhibitory diode can thereafter strengthen.

At all times during the learning process, a weak alternating electric field can be applied perpendicular to the connections. This can cause the connections to weaken by rotating the nanotube perpendicular to the connection direction. This perpendicular field is important because it can allow for a much higher degree of adaptation. To understand this, one must realize that the connections cannot (practically) keep getting stronger and stronger. By weakening those connections not contributing much to the desired output, we decrease the necessary strength of the needed connections and allow for more flexibility in continuous training. This perpendicular alternating voltage can be realized by the addition of two electrodes on the outer extremity of the connection set, such as plates sandwiching the connections (i.e., above and below). Other mechanisms, such as increasing the temperature of the nanotube suspension could also be used for such a purpose, although this method is perhaps a little less controllable or practical.

The circuit depicted in FIG. **7** can be separated into two separate circuits. The first part of the circuit can be composed of nanotube connections, while the second part of the circuit comprises the "neurons" and the learning mechanism (i.e., op-amps/comparator). The learning mechanism on first glance appears similar to a relatively standard circuit that could be implemented on silicon with current technology. Such a silicon implementation can thus comprise the "neuron" chip. The second part of the circuit (i.e., the connections) is thus a new type of chip, although it could be constructed with current technology. The connection chip can be composed of an orderly array of electrodes spaced anywhere from, for example, 100 nm to 1 μm or perhaps even further. In a biological system, one talks of synapses connecting neurons. It is in the synapses where the information is processed, (i.e., the "connection weights"). Similarly, such a chip can contain all of the synapses for the physical neural network. A possible arrangement thereof can be seen in FIG. **8**.

FIG. **8** illustrates a schematic diagram of a chip layout **800** for a connection network that may be implemented in accordance with an alternative embodiment. FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., connection network **800**) that can be implemented in accordance with one or more embodiments. Chip layout **800** includes an input array composed of plurality of inputs **801**, **802**, **803**, **804**, and **805**, which are provided to a plurality of layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. A plurality of outputs **802** can be derived from layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. Inputs **801** can be coupled to layers **806** and **807**, while inputs **802** can be connected to layers **808** and **809**. Similarly, inputs **803** can be connected to layers **810** and **811**. Also, inputs **804** can be connected to layers **812** and **813**. Inputs **805** are generally connected to layers **814** and **815**.

Similarly, such an input array can includes a plurality of inputs **831**, **832**, **833**, **834** and **835** which are respectively input to a plurality of layers **816**, **817**, **818**, **819**, **820**, **821**, **822**, **823**, **824** and **825**. Thus, inputs **831** can be connected to layers **816** and **817**, while inputs **832** are generally coupled to layers **818** and **819**. Additionally, inputs **833** can be connected to layers **820** and **821**. Inputs **834** can be connected to layers **822** and **823**. Finally, inputs **835** are connected to layers **824** and **825**. Arrows **828** and **830** represent a continuation of the aforementioned connection network pattern. Those skilled in the art can appreciate, of course, that chip layout **800** is not intended to represent an exhaustive chip layout or to limit the scope of the invention. Many modifications and variations to chip layout **800** are possible in light of the teachings herein without departing from the scope of the embodiments. It is contemplated that the use of a chip layout, such as chip layout **800**, can involve a variety of components having different characteristics.

Preliminary calculations based on a maximum etching capability of 200 nm resolution indicated that over 4 million synapses could fit on an area of approximately 1 cm^2. The smallest width that an electrode can possess is generally based on current lithography. Such a width may of course change as the lithographic arts advance. This value is actually about 70 nm for state-of-the-art techniques currently. These calculations are of course extremely conservative, and are not considered a limiting feature of the embodiments. Such calculations are based on an electrode

with, separation, and gap of approximately 200 nm. For such a calculation, for example, 166 connection networks comprising 250 inputs and 100 outputs can fit within a one square centimeter area.

If such chips are stacked vertically, an untold number of synapses could be attained. This is two to three orders of magnitude greater than some of the most capable neural network chips out there today, chips that rely on standard methods to calculate synapse weights. Of course, the geometry of the chip could take on many different forms, and it is quite possible (based on a conservative lithography and chip layout) that many more synapses could fit in the same space. The training of a chip this size would take a fraction of the time of a comparably sized traditional chip using digital technology.

The training of such a chip is primarily based on two assumptions. First, the inherent parallelism of a physical neural network (i.e., a Knowm) can permit all training sessions to occur simultaneously, no matter now large the associated connection network. Second, recent research has indicated that near perfect aligning of nanotubes can be accomplished in approximately 15 minutes. If one considers that the input data, arranged as a vector of binary "high's" and "low's" is presented to the Knowm simultaneously, and that all training vectors are presented one after the other in rapid succession (e.g., perhaps 100 MHz or more), then each connection would "see" a different frequency in direct proportion to the amount of time that its connection is required for accurate data processing (i.e., provided by a feedback mechanism). Thus, if it only takes approximately 15 minutes to attain an almost perfect state of alignment, then this amount of time would comprise the longest amount of time required to train, assuming that all of the training vectors are presented during that particular time period.

FIG. **9** illustrates a flow chart **900** of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment. Initially, as indicated at block **902**, a connection gap is created from a connection network structures. As indicated earlier, the goal for such a connection network is generally to develop a network of connections of "just" the right values to satisfy particular information processing requirements, which is precisely what a neural network accomplishes. As illustrated at block **904**, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent," and also a suspension.

The solvent utilized can comprise a volatile liquid that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution may be sandwiched between wafers of silicon or other materials. If the fluid has a melting point that is approximately at room temperature, then the viscosity of the fluid could be controlled easily. Thus, if it is desired to lock the connection values into a particular state, the associated physical neural network (i.e., Knowm) may be cooled slightly until the fluid freezes. The term "solvent" as utilized herein thus can include fluids such as for example, toluene, hexadecane, mineral oil, etc. Note that the solution in which the nanoconductors (i.e., nanoconnections) are present should generally comprise a dielectric. Thus, when the resistance between the electrodes is measured, the conductivity of the nanoconductors can be essentially measured, not that of the solvent. The nanoconductors can be suspended in the solution or can alternately lie on the bottom surface of the connection gap. The solvent may also be provided in the form of a gas.

As illustrated thereafter at block **906**, the nanoconductors must be suspended in the solvent, either dissolved or in a suspension of sorts, but generally free to move around, either in the solution or on the bottom surface of the gap. As depicted next at block **908**, the electrical conductance of the solution must be less than the electrical conductance of the suspended nanoconductor(s). Similarly, the electrical resistance of the solution is greater than the electrical resistance of the nanoconductor.

Next, as illustrated at block **910**, the viscosity of the substance should not be too much so that the nanoconductors cannot move when an electric field (e.g., voltage) is applied. Finally, as depicted at block **912**, the resulting solution of the "solvent" and the nanoconductors is thus located within the connection gap.

Note that although a logical series of steps is illustrated in FIG. **9**, it can be appreciated that the particular flow of steps can be re-arranged. Thus, for example, the creation of the connection gap, as illustrated at block **902**, may occur after the preparation of the solution of the solvent and nanoconductor(s), as indicated at block **904**. FIG. **9** thus represents merely possible series of steps, which may be followed to create a connection network. A variety of other steps may be followed as long as the goal of achieving a connection network is achieved. Similar reasoning also applies to FIG. **10**.

FIG. **10** illustrates a flow chart **1000** of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment. As indicated at block **1002**, an electric field can be applied across the connection gap discussed above with respect to FIG. **9**. The connection gap can be occupied by the solution discussed above. As indicated thereafter at block **1004**, to create the connection network, the input terminals can be selectively raised to a positive voltage while the output terminals are selectively grounded. As illustrated thereafter at block **1006**, connections thus form between the inputs and the outputs. The important requirements that make the resulting physical neural network functional as a neural network is that the longer this electric field is applied across the connection gap, or the greater the frequency or amplitude, the more nanoconductors align and the stronger the connection becomes. Thus, the connections that get utilized the most frequently become the strongest.

As indicated at block **1008**, the connections can either be initially formed and have random resistances or no connections will be formed at all. By forming initial random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up as much. Depending on the rate of connection decay, having initial random connections could prove to be a faster method, although not necessarily. A connection network can adapt itself to whatever is required regardless of the initial state of the connections. Thus, as indicated at block **1010**, as the electric field is applied across the connection gap, the more the nonconductor(s) will align and the stronger the connection becomes. Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block **1012**. As illustrated at block **1014**, the resistance of the connection can be maintained or lowered by selective activations of the connections. In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain.

The neurons in a human brain, although seemingly simple when viewed individually, interact in a complicated network that computes with both space and time. The most basic picture of a neuron, which is usually implemented in technology, is a summing device that adds up a signal. Actually, this statement can be made even more general by stating that a neuron adds up a signal in discrete units of time. In other words, every group of signals incident upon the neuron can be viewed as occurring in one moment in time. Summation thus occurs in a spatial manner. The only difference between one signal and another signal depends on where such signals originate. Unfortunately, this type of data processing excludes a large range of dynamic, varying situations that cannot necessarily be broken up into discrete units of time.

The example of speech recognition is a case in point. Speech occurs in the time domain. A word is understood as the temporal pronunciation of various syllables. A sentence is composed of the temporal separation of varying words. Thoughts are composed of the temporal separation of varying sentences. Thus, for an individual to understand a spoken language at all, a syllable, word, sentence or thought must exert some type of influence on another syllable, word, sentence or thought. The most natural way that one sentence can exert any influence on another sentence, in the light of neural networks, is by a form of temporal summation. That is, a neuron "remembers" the signals it received in the past.

The human brain accomplishes this feat in an almost trivial manner. When a signal reaches a neuron, the neuron has an influx of ions rush through its membrane. The influx of ions contributes to an overall increase in the electrical potential of the neuron. Activation is achieved when the potential inside the cell reaches a certain threshold. The one caveat is that it takes time for the cell to pump out the ions, something that it does at a more or less constant rate. So, if another signal arrives before the neuron has time to pump out all of the ions, the second signal will add with the remnants of the first signal and achieve a raised potential greater than that which could have occurred with only the second signal. The first signal influences the second signal, which results in temporal summation.

Implementing this in a technological manner has proved difficult in the past. Any simulation would have to include a "memory" for the neuron. In a digital representation, this requires data to be stored for every neuron, and this memory would have to be accessed continually. In a computer simulation, one must discritize the incoming data, since operations (such as summations and learning) occur serially. That is, a computer can only do one thing at a time. Transformations of a signal from the time domain into the spatial domain require that time be broken up into discrete lengths, something that is not necessarily possible with real-time analog signals in which no point exists within a time-varying signal that is uninfluenced by another point.

A physical neural network, however, is generally not digital. A physical neural network is a massively parallel analog device. The fact that actual molecules (e.g., nanoconductors) must move around (in time) makes temporal summation a natural occurrence. This temporal summation is built into the nanoconnections. The easiest way to understand this is to view the multiplicity of nanoconnections as one connection with one input into a neuron-like node (Op-amp, Comparator, etc.). This can be seen in FIG. **11**.

FIG. **11** illustrates a schematic diagram of a circuit **1100** illustrating temporal summation within a neuron, in accordance with a preferred embodiment. As indicated in FIG. **11**, an input **1102** can be provided to nanoconnections **1104**, which in turn can provide a signal, which can be input to an amplifier **1110** (e.g., op amp) at node B. A resistor **1106** can be connected to node A, which in turn is electrically equivalent to node B. Node B can be connected to a negative input of amplifier **1100**. Resistor **1108** can also be connected to a ground **1108**. Amplifier **1110** provides output **1114**. Note that although nanoconnections **1104** is referred to in the plural it can be appreciated that nanoconnections **1104** can comprise a single nanoconnection or a plurality of nanoconnections. For simplicity sake, however, the plural form is used to refer to nanoconnections **1104**.

Input **1102** can be provided by another physical neural network (i.e., Knowm) to cause increased connection strength of nanoconnections **1104** over time. This input would most likely arrive in pulses, but could also be continuous. A constant or pulsed electric field perpendicular to the connections can serve to constantly erode the connections, so that only signals of a desired length or amplitude can cause a connection to form. Once the connection is formed, the voltage divider formed by nanoconnection **1104** and resistor **1106** can cause a voltage at node A in direct proportion to the strength of nanoconnections **1104**. When the voltage at node A reaches a desired threshold, the amplifier (i.e., an op-amp and/or comparator), will output a high voltage (i.e., output **1114**). The key to the temporal summation is that, just like a real neuron, it takes time for the electric field to breakdown the nanoconnections **1104**, so that signals arriving close in time will contribute to the firing of the neuron (i.e., op-amp, comparator, etc.). Temporal summation has thus been achieved. The parameters of the temporal summation could be adjusted by the amplitude and frequency of the input signals and the perpendicular electric field.

FIG. **12** illustrates a block diagram illustrating a pattern recognition system **1200**, which may be implemented with a physical neural network device **1222**, in accordance with an alternative embodiment. Note that pattern recognition system **1200** can be implemented as a speech recognition system. Although pattern recognition system **1200** is depicted herein in the context of speech recognition, a physical neural network device (i.e., a Knowm device) may be implemented with other pattern recognition systems, such as visual and/or imaging recognition systems. FIG. **12** thus does not comprise a limiting feature of the embodiments and is presented for general edification and illustrative purposes only. Those skilled in the art can appreciate that the diagram depicted in FIG. **12** may be modified as new applications and hardware are developed. The development or use of a pattern recognition system such as pattern recognition system **1200** of FIG. **12** by no means limits the scope of the physical neural network (i.e., Knowm) disclosed herein.

FIG. **12** thus illustrates in block diagram fashion, the system structure of a speech recognition device using a neural network according to an alternative embodiment. The pattern recognition system **1200** can be provided with a CPU **1211** for performing the functions of inputting vector rows and instructor signals (vector rows) to an output layer for the learning process of a physical neural network device **1222**, and changing connection weights between respective neuron devices based on the learning process. Pattern recognition system **1200** can be implemented within the context of a data-processing system, such as, for example, a personal computer or personal digital assistant (PDA), both of which are well known in the art.

The CPU **1211** can perform various processing and controlling functions, such as pattern recognition, including but not limited to speech and/or visual recognition based on the output signals from the physical neural network device

1222. The CPU **1211** is connected to a read-only memory (ROM) **1213**, a random-access memory (RAM) **1214**, a communication control unit **1215**, a printer **1216**, a display unit **1217**, a keyboard **1218**, an FFT (fast Fourier transform) unit **1221**, a physical neural network device **1222** and a graphic reading unit **1224** through a bus line **1220** such as a data bus line. The bus line **1220** may comprise, for example, an ISA, EISA, or PCI bus.

The ROM **1213** is a read-only memory storing various programs or data used by the CPU **1211** for performing processing or controlling the learning process, and speech recognition of the physical neural network device **1222**. The ROM **1213** may store programs for carrying out the learning process according to error back-propagation for the physical neural network device or code rows concerning, for example, 80 kinds of phonemes for performing speech recognition. The code rows concerning the phonemes can be utilized as second instructor signals and for recognizing phonemes from output signals of the neuron device network. Also, the ROM **1213** can store programs of a transformation system for recognizing speech from recognized phonemes and transforming the recognized speech into a writing (i.e., written form) represented by characters.

A predetermined program stored in the ROM **1213** can be downloaded and stored in the RAM **1214**. RAM **1214** generally functions as a random access memory used as a working memory of the CPU **1211**. In the RAM **1214**, a vector row storing area can be provided for temporarily storing a power obtained at each point in time for each frequency of the speech signal analyzed by the FFT unit **1221**. A value of the power for each frequency serves as a vector row input to a first input portion of the physical neural network device **1222**. Further, in the case where characters or graphics are recognized in the physical neural network device, the image data read by the graphic reading unit **1224** are stored in the RAM **1214**.

The communication control unit **1215** transmits and/or receives various data such as recognized speech data to and/or from another communication control unit through a communication network **1202** such as a telephone line network, an ISDN line, a LAN, or a personal computer communication network. Network **1202** may also comprise, for example, a telecommunications network, such as a wireless communications network. Communication hardware methods and systems thereof are well known in the art.

The printer **1216** can be provided with a laser printer, a bubble-type printer, a dot matrix printer, or the like, and prints contents of input data or the recognized speech. The display unit **1217** includes an image display portion such as a CRT display or a liquid crystal display, and a display control portion. The display unit **1217** can display the contents of the input data or the recognized speech as well as a direction of an operation required for speech recognition utilizing a graphical user interface (GUI).

The keyboard **1218** generally functions as an input unit for varying operating parameters or inputting setting conditions of the FFT unit **1221**, or for inputting sentences. The keyboard **1218** is generally provided with a ten-key numeric pad for inputting numerical figures, character keys for inputting characters, and function keys for performing various functions. A mouse **1219** can be connected to the keyboard **1218** and serves as a pointing device.

A speech input unit **1223**, such as a microphone can be connected to the FFT unit **1221**. The FFT unit **1221** transforms analog speech data input from the voice input unit **1223** into digital data and carries out spectral analysis of the digital data by discrete Fourier transformation. By performing a spectral analysis using the FFT unit **1221**, the vector row based on the powers of the respective frequencies are output at predetermined intervals of time. The FFT unit **1221** performs an analysis of time-series vector rows, which represent characteristics of the inputted speech. The vector rows output by the FFT **1221** are stored in the vector row storing area in the RAM **1214**.

The graphic reading unit **224**, provided with devices such as a CCD (Charged Coupled Device), can be used for reading images such as characters or graphics recorded on paper or the like. The image data read by the image-reading unit **1224** are stored in the RAM **1214**. Note that an example of a pattern recognition apparatus, which may be modified for use with the physical neural network described herein, is disclosed in U.S. Pat. No. 6,026,358 to Tomabechi, Feb. 16, 2000, "Neural Network, A Method of Learning of a Neural Network and Phoneme Recognition Apparatus Utilizing a Neural Network." U.S. Pat. No. 6,026,358 is incorporated herein by reference. It can be appreciated that the Tomabechi reference does not teach, suggest or anticipate the embodiments, but is discussed herein for general illustrative, background and general edification purposes only.

The implications of a physical neural network are tremendous. With existing lithography technology, many electrodes in an array such as depicted in FIG. **5** can be etched onto a wafer of silicon. The neuron-diodes, as well as the training circuitry illustrated in FIG. **6**, could be built onto the same silicon wafer, although it may be desirable to have the connections on a separate chip due to the liquid solution of nanoconductors. A solution of suspended nanoconductors could be placed between the electrode connections and the chip could be packaged. The resulting "chip" would look much like a current Integrated Chip (IC) or VLSI (very large scale integrated) chips. One could also place a rather large network parallel with a computer processor as part of a larger system. Such a network, or group of networks, could add significant computational capabilities to standard computers and associated interfaces.

For example, such a chip may be constructed utilizing a standard computer processor in parallel with a large physical neural network or group of physical neural networks. A program can then be written such that the standard computer teaches the neural network to read, or create an association between words, which is precisely the same sort of task in which neural networks can be implemented. Once the physical neural network is able to read, it can be taught for example to "surf" the Internet and find material of any particular nature. A search engine can then be developed that does not search the Internet by "keywords", but instead by meaning. This idea of an intelligent search engine has already been proposed for standard neural networks, but until now has been impractical because the network required was too big for a standard computer to simulate. The use of a physical neural network (i.e., physical neural network) as disclosed herein now makes a truly intelligent search engine possible.

A physical neural network can be utilized in other applications, such as, for example, speech recognition and synthesis, visual and image identification, management of distributed systems, self-driving cars, filtering, etc. Such applications have to some extent already been accomplished with standard neural networks, but are generally limited in expense, practicality and not very adaptable once implemented. The use of a physical neural network can permit such applications to become more powerful and adaptable. Indeed, anything that requires a bit more "intelligence" could incorporate a physical neural network. One of the

primary advantages of a physical neural network is that such a device and applications thereof can be very inexpensive to manufacture, even with present technology. The lithographic techniques required for fabricating the electrodes and channels therebetween has already been perfected and implemented in industry.

Most problems in which a neural network solution is implemented are complex adaptive problems, which change in time. An example is weather prediction. The usefulness of a physical neural network is that it could handle the enormous network needed for such computations and adapt itself in real-time. An example wherein a physical neural network (i.e., Knowm) can be particularly useful is the Personal Digital Assistant (PDA). PDA's are well known in the art. A physical neural network applied to a PDA device can be advantageous because the physical neural network can ideally function with a large network that could constantly adapt itself to the individual user without devouring too much computational time from the PDA. A physical neural network could also be implemented in many industrial applications, such as developing a real-time systems control to the manufacture of various components. This systems control can be adaptable and totally tailored to the particular application, as necessarily it must.

It will be appreciated that variations of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also that various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

What is claimed is:

1. A temporal summation device, comprising:

at least one nanoconnection having an input and an output thereof, wherein an input signal provided to said input causes said at least one nanoconnection to experience an increase in connection strength thereof over time;

a voltage divider formed by said at least one nanoconnection and a resistor connected to said output of said at least one nanoconnection, wherein said voltage divider provides a voltage at said output of said at least one nanoconnection that is in direct proportion to said connection strength of said at least one nanoconnection; and

an amplifier connected to said voltage divider, wherein when said voltage provided by said voltage divider attains a desired threshold voltage, said amplifier attains a high voltage output thereby providing a temporal summation device thereof.

2. The device of claim **1** wherein an electric field perpendicular to said at least one nanoconnection serves to constantly erode said at least one nanoconnection so that only desired signals cause said at least one nanoconnection to form.

3. The device of claim **2** wherein said desired signals comprise signals of a desired length.

4. The device of claim **2** wherein said desired signals comprise signals of a desired amplitude.

5. The device of claim **2** where said electric field perpendicular to said at least one nanoconnection comprises a constant electric field.

6. The device of claim **2** wherein said electric field perpendicular to said at least one nanoconnection comprises a pulsed electric field.

7. The device of claim **2** wherein said at least one parameter of said temporal summation device is further adjustable by varying said electric field.

8. The device of claim **1** wherein said amplifier comprises an operational amplifier.

9. The device of claim **1** wherein said amplifier comprises a comparator.

10. The device of claim **1** wherein said amplifier comprises a negative input, a positive input and an output thereof, wherein said negative input is connected to said resistor and said output of said at least one nanoconnection and wherein said positive input of said amplifier is connected to said output of said amplifier, such that said input signal provided at said input of said at least one nanoconnection generates an output signal at said output of said at least one nanoconnection at said first node that is in turn input to said amplifier at said negative input of said amplifier.

11. The device of claim **10** wherein said desired threshold voltage is identifiable at said positive input of said amplifier.

12. The device of claim **1** wherein said resistor is connected to ground and said negative input of said amplifier at said output of said at least one nanoconnection.

13. The device of claim **1** wherein said amplifier comprises an artificial neuron.

14. The device of claim **1** wherein at least one parameter of said temporal summation device is adjustable by varying an amplitude and a frequency of at least one input signal input to said at least one nanoconnection.

15. A temporal summation device, comprising:

at least one nanoconnection having an input and an output thereof, wherein an input signal provided to said input causes said at least one nanoconnection to experience an increase in connection strength thereof over time and wherein an electric field perpendicular to said at least one nanoconnection serves to constantly erode said at least one nanoconnection so that only desired signals cause said at least one nanoconnection to form;

a voltage divider formed by said at least one nanoconnection and a resistor connected to said output of said at least one nanoconnection, wherein said voltage divider provides a voltage at said output of said at least one nanoconnection that is in direct proportion to said connection strength of said at least one nanoconnection, wherein said resistor is connected to ground and said negative input of said amplifier at said output of said at least one nanoconnection; and

an amplifier connected to said voltage divider, wherein when said voltage provided by said voltage divider attains a desired threshold voltage identifiable at said positive input of said amplifier, said amplifier attains a high voltage output thereby providing a temporal summation device thereof, wherein at least one parameter of said temporal device is adjustable by varying an amplitude and a frequency of at least one input signal input to said at least one nanoconnection and/or said electric field perpendicular to said at least one nanoconnection.

16. The device of claim **15** where said electric field perpendicular to said at least one nanoconnection comprises a constant electric field.

17. The device of claim **15** wherein said electric field perpendicular to said at least one nanoconnection comprises a pulsed electric field.

18. A temporal summation device, comprising:

at least one nanoconnection having an input and an output thereof, wherein an input signal provided to said input

causes said at least one nanoconnection to experience an increase in connection strength thereof over time and wherein an electric field perpendicular to said at least one nanoconnection serves to constantly erode said at least one nanoconnection so that only desired signals cause said at least one nanoconnection to form;

a voltage divider formed by said at least one nanoconnection and a resistor connected to said output of said at least one nanoconnection, wherein said voltage divider provides a voltage at said output of said at least one nanoconnection that is in direct proportion to said connection strength of said at least one nanoconnection, wherein said resistor is connected to ground and said negative input of said amplifier at said output of said at least one nanoconnection; and

an amplifier connected to said voltage divider, wherein when said voltage provided by said voltage divider attains a desired threshold voltage identifiable at said positive input of said amplifier, said amplifier attains a high voltage output thereby providing a temporal summation device thereof, wherein at least one parameter of said temporal device is adjustable by varying an amplitude and a frequency of at least one input signal input to said at least one nanoconnection and/or said electric field perpendicular to said at least one nanoconnection,

wherein said amplifier comprises a negative input, a positive input and an output thereof, wherein said negative input is connected to said resistor and said output of said at least one nanoconnection and wherein said positive input of said amplifier is connected to said output of said amplifier, such that said input signal provided at said input of said at least one nanoconnection generates an output signal at said output of said at least one nanoconnection at said first node that is in turn input to said amplifier at said negative input of said amplifier.

19. The device of claim **18** where said electric field perpendicular to said at least one nanoconnection comprises a constant electric field or a pulsed electric field.

20. The device of claim **18** wherein at least one parameter of said temporal summation device is adjustable by varying an amplitude and a frequency of at least one input signal input to said at least one nanoconnection and/or by varying said electric field.

* * * * *

US006995649B2

United States Patent

(12) **United States Patent**
Nugent

(10) **Patent No.: US 6,995,649 B2**
(45) **Date of Patent: Feb. 7, 2006**

(54) **VARIABLE RESISTOR APPARATUS FORMED UTILIZING NANOTECHNOLOGY**

(75) Inventor: **Alex Nugent**, Santa Fe, NM (US)

(73) Assignee: **KnowmTech, LLC**, Albuquerque, NM (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **11/045,846**

(22) Filed: **Jan. 31, 2005**

(65) **Prior Publication Data**

US 2005/0151615 A1 Jul. 14, 2005

Related U.S. Application Data

(63) Continuation of application No. 10/095,273, filed on Mar. 12, 2002, now Pat. No. 6,889,216.

(51) **Int. Cl.**
H01C 7/10 (2006.01)

(52) **U.S. Cl.** **338/20**; 338/21

(58) **Field of Classification Search** 338/20, 338/21

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

2,707,223 A *	4/1955	Hollmann	338/32 R
3,833,894 A *	9/1974	Aviram et al.	365/151
4,802,951 A	2/1989	Clark et al.	156/630
4,974,146 A	11/1990	Works et al.	364/200
4,988,891 A	1/1991	Mashiko	307/201
5,315,162 A	5/1994	McHardy et al.	307/201
5,422,983 A	6/1995	Castelaz et al.	395/24
5,475,794 A	12/1995	Mashiko	395/24
5,589,692 A	12/1996	Reed	257/23
5,649,063 A	7/1997	Bose	395/22
5,670,818 A *	9/1997	Forouhi et al.	257/530
5,706,404 A	1/1998	Colak	395/24
5,717,832 A	2/1998	Steimle et al.	395/24
5,761,115 A *	6/1998	Kozicki et al.	365/182
5,783,840 A	7/1998	Randall et al.	257/24
5,812,993 A	9/1998	Ginosar et al.	706/26
5,896,312 A *	4/1999	Kozicki et al.	365/153
5,904,545 A	5/1999	Smith et al.	438/455
5,914,893 A *	6/1999	Kozicki et al.	365/107

(Continued)

FOREIGN PATENT DOCUMENTS

EP 1 022 764 A1 1/2000

(Continued)

OTHER PUBLICATIONS

"Nanoparticles Get Wired", Dimes Institute, Delft University of Technology, 1997.*

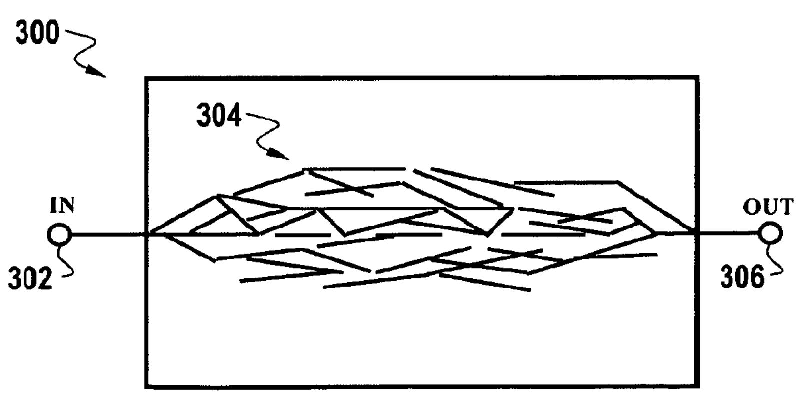
(Continued)

Primary Examiner—George Davis
(74) *Attorney, Agent, or Firm*—Kermit O. Lopez; Luis M. Ortiz; Ortiz & Lopez, PLLC

(57) **ABSTRACT**

A variable resistor apparatus includes a plurality of nanoparticles disposed between two terminals, wherein the plurality of nanoparticles provides an electrical resistance. An electric field applied to the plurality of nanoparticles across the two terminals results in an alignment of the nanoparticles over time and a decrease in the electrical resistance thereby providing a variable resistor apparatus. The electric or electrical field can be applied across the two terminals perpendicular to the plurality of nanoconnections. The nanoparticles can comprise nanoconductors, which can be formed as, for example, nanotubes and/or nanowires. The nanoparticles are generally disposed in a solution within a connection gap formed between the two terminals. The solution can comprise a solvent and/or a suspension of nanoparticles forming a mixture thereof. The solution can also be provided as a liquid, a gel, and or a gas. The solution may also comprise a dielectric.

19 Claims, 6 Drawing Sheets



U.S. PATENT DOCUMENTS

5,951,881	A	9/1999	Rogers et al.	216/41
5,978,782	A	11/1999	Neely	706/16
6,026,358	A	2/2000	Tomabechi	704/232
6,084,796	A *	7/2000	Kozicki et al.	365/153
6,128,214	A	10/2000	Kuekes et al.	365/151
6,245,630	B1	6/2001	Ishikawa	438/393
6,248,529	B1	6/2001	Connolly	435/6
6,256,767	B1	7/2001	Kuekes et al.	716/9
6,282,530	B1	8/2001	Huang	706/41
6,294,450	B1	9/2001	Chen et al.	438/597
6,314,019	B1	11/2001	Kuekes et al.	365/151
6,330,553	B1	12/2001	Uchikawa et al.	706/2
6,335,291	B1	1/2002	Freeman	438/706
6,339,227	B1	1/2002	Ellenbogen	257/40
6,359,288	B1	3/2002	Ying et al.	257/14
6,363,369	B1	3/2002	Liaw et al.	706/15
6,383,923	B1	5/2002	Brown et al.	438/666
6,389,404	B1	5/2002	Carson et al.	706/18
6,407,443	B2	6/2002	Chen et al.	257/616
6,418,423	B1	7/2002	Kambhatla et al.	706/15
6,420,092	B1	7/2002	Yang et al.	430/311
6,422,450	B1	7/2002	Zhou et al.	228/121.85
6,423,583	B1	7/2002	Avouris et al.	438/132
6,424,961	B1	7/2002	Ayala	706/25
6,426,134	B1	7/2002	Lavin et al.	428/300.1
6,620,346	B1 *	9/2003	Schulz et al.	252/519.51
6,798,692	B2 *	9/2004	Kozicki et al.	365/174
2001/0004471	A1	6/2001	Zhang	427/372.2
2001/0023986	A1	9/2001	Mancevski	257/741
2001/0024633	A1	9/2001	Lee et al.	423/447.3
2001/0031900	A1 *	10/2001	Margrave et al.	570/126
2001/0041160	A1 *	11/2001	Margrave et al.	423/460
2001/0044114	A1	11/2001	Connolly	435/6
2002/0001905	A1	1/2002	Choi et al.	438/268
2002/0004028	A1 *	1/2002	Margrave et al.	423/447.3
2002/0004136	A1	1/2002	Gao et al.	428/367
2002/0030205	A1	3/2002	Varshavsky	257/208
2002/0075126	A1 *	6/2002	Reitz et al.	338/21
2002/0086124	A1 *	7/2002	Margrave et al.	428/36.9
2002/0090468	A1	7/2002	Goto et al.	427/580
2002/0102353	A1	8/2002	Mauthner et al.	427/255.28
2003/0031438	A1 *	2/2003	Kambe et al.	385/122
2003/0177450	A1	9/2003	Nugent	716/1
2003/0236760	A1	12/2003	Nugent	706/26
2004/0039717	A1	2/2004	Nugent	706/27
2004/0150010	A1	8/2004	Snider	257/209
2004/0153426	A1	8/2004	Nugent	706/25
2004/0162796	A1	8/2004	Nugent	706/27
2004/0193558	A1	9/2004	Nugent	706/25

FOREIGN PATENT DOCUMENTS

EP	1 046 613	A2	4/2000
EP	1 100 106	A2	5/2001
EP	1 069 206	A2	7/2001
EP	1 115 135	A1	7/2001
EP	1 134 304	A2	9/2001
RU	2071126	C1 *	6/1996
WO	WO 00/44094		7/2000
WO	WO 03/017282	A1 *	8/2001

OTHER PUBLICATIONS

A. Bezryadin, "Trapping Single Particle with Nanoelectrodes", Physics News Graphics, Sep. 1997.*

Snow et al, "Nanofabrication with Proximal Probes", Proceedings of the IEEE, Apr. 1997.*

Peter Weiss, "Circuitry in a Nanowire: Novel Growth Method May Transform Chips," Science News Online, vol. 161, No. 6; Feb. 9, 2002.

Press Release, "Nanowire-based electronics and optics comes one step closer," Eureka Alert, American Chemical Society; Feb. 1, 2002.

Weeks et al., "High-pressure nanolithography using low-energy electrons from a scanning tunneling microscope," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 38-42; Dec. 12, 2001.

CMP Cientifica, "Nanotech: the tiny revolution"; CMP Cientifica, Nov. 2001.

Diehl, et al., "Self-Assembled, Deterministic Carbon Nanotube Wiring Networks," Angew. Chem. Int. Ed. 2002, 41, No. 2; Received Oct. 22, 2001.

G. Pirio, et al., "Fabrication and electrical characteristics of carbon nanotube field emission microcathodes with an integrated gate electrode," Insititute of Physics Publishing, Nanotechnology 13 (2002), pp. 1-4, Oct. 2, 2001.

Leslie Smith, "An Introduction to Neural Networks," Center for Cognitive and Computational Neuroscience, Dept. of Computing & Mathematics, University of Stirling, Oct. 25, 1996; http//www.cs.stir.ac.uk/~lss/NNIntro/InvSlides.html.

V. Derycke et al., "Carbon Nanotube Inter- and Intramolecular Logic Gates," American Chemical Society, Nano Letters, XXXX, vol. 0, No. 0, A-D.

Mark K. Anderson, "Mega Steps Toward the Nanochip," Wired News, Apr. 27, 2001.

Collins et al., "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," Science, vol. 292, pp. 706-709, Apr. 27, 2001.

Landman et al., "Metal-Semiconductor Nanocontacts: Silicon Nanowires," Physical Review Letters, vol. 85, No. 9, Aug. 28, 2000.

John G. Spooner, "Tiny tubes mean big chip advances," Cnet News.com, Tech News First, Apr. 26, 2001.

Jeong-Mi Moon et al., "High-Yield Purification Process of Singlewalled Carbon Nanotubes," J. Phys. Chem. B 2001, 105, pp. 5677-5681.

"A New Class of Nanostructure: Semiconducting Nanobelts Offer Potential for Nanosensors and Nanoelectronics," Mar. 12, 2001, http://www.sciencedaily.com/releases/2001/03/010309080953.htm.

Hermanson et al., "Dielectrophoretic Assembly of Electrically Functional Microwires from Nanoparticle Suspensions," Materials Science, vol. 294, No. 5544, Issue of Nov. 2, 2001, pp. 1082-1086.

Press Release, "Toshiba Demonstrates Operation of Single-Electron Transistor Circuit at Room Temperature," Toshiba, Jan. 10, 2001.

J. Appenzeller et al., "Optimized contact configuration for the study of transport phenomena in ropes of single-wall carbon nanotubes," Applied Physics Letters, vol. 78, No. 21, pp. 3313-3315, May 21, 2001.

David Rotman, "Molecular Memory, Replacing silicon with organic molecules could mean tiny supercomputers," Technology Review, May 2001, p. 46.

Westervelt et al., "Molecular Electronics," NSF Functional Nanostructures Grant 9871810, NSF Partnership in Nanotechnology Conference, Jan. 29-30, 2001; http://www.unix.oit.umass.edu/~nano/NewFiles/FN19_Harvard.pdf.

Nyogi et al., "Chromatographic Purification of Soluble Single-Walled Carbon Nanotubes (s-SWNTs)," J. Am. Chem. Soc 2001, 123, pp. 733-734, Received Jul. 10, 2000.

Duan et al., "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices," Nature, vol. 409, Jan. 4, 2001, pp. 67-69.

Paulson, et al., "Tunable Resistance of a Carbon Nanotube-Graphite Interface," Science, vol. 290, Dec. 1, 2000, pp. 1742-1744.
Wei et al., "Reliability and current carrying capacity of carbon nanotubes," Applied Physics Letters, vol. 79, No. 8, Aug. 20, 2001, pp. 1172-1174.
Collins et al., "Nanotubes for Electronics," Scientific American, Dec. 2000, pp. 62-69.
Avouris et al., "Carbon nanotubes: nanomechanics, manipulation, and electronic devices," Applied Surface Science 141 (1999), pp. 201-209.
Smith et al., "Electric-field assisted assembly and alignment of metallic nanowires," Applied Physics Letters, vol. 77, No. 9, Aug. 28, 2000, pp. 1399-1401.
Hone et al., "Electrical and thermal transport propertes of mangetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 666-668.
Smith et al., "Structural anisotropy of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 663-665.
Andriotis et al., "Various bonding configurations of transition-metal atoms on carbon nanotubes: Their effect on contact resistance," Applied Physics Letters, vol. 76, No. 26, Jun. 26, 2000, pp. 3890-3892.
Chen et al., "Aligning single-wall carbon nanotubes with an alternating-current electric field," Applied Physics Letters, vol. 78, No. 23, Jun. 4, 2001, pp. 3714-3716.
Bezryadin et al., "Self-assembled chains of graphitized carbon nanoparticles," Applied Physics Letters, vol. 74, No. 18, May 3, 1999, pp. 2699-2701.
Bezryadin et al., "Evolution of avalanche conducting states in electrorheological liquids," Physical Review E, vol. 59, No. 6, Jun. 1999, pp. 6896-6901.
Liu et al., "Fullerene Pipes," Science, vol. 280, May 22, 1998, pp. 1253-1255.
Yamamoto et al., "Orientation and purification of carbon nanotubes using ac electrophoresis," J. Phys. D: Appl. Phys 31 (1998) L34-L36.
Bandow et al., "Purification of Single-Wall Carbon Nanotubes by Microfiltration," J. Phys. Chem. B 1997, 101, pp. 8839-8842.
Tohji et al., "Purifying single walled nanotubes," Nature, vol. 383, Oct. 24, 1996, p. 679.
Dejan Rakovic, "Hierarchical Neural Networks and Brainwaves: Towards a Theory of Consciousness," Brain & Consciousness: Proc. ECPD Workshop (ECPD), Belgrade, 1997), pp. 189-204.
Dave Anderson & George McNeill, "Artificial Neural Networks Technology," A DACS (Data & Analysis Center for Software) State-of-the-Art Report, Contract No. F30602-89-C-0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, Aug. 20, 1992.
Greg Mitchell, "Sub-50 nm Device Fabrication Stragegies," Project No. 890-00, Cornell Nanofabrication Facility, Electronics—p. 90-91, National Nanofabrication Users Network.
John-William DeClaris, "An Introduction to Neural Networks," http://www.ee.umd.edu/medlab/neural/nn1.html.
"Neural Networks," StatSoft, Inc., http://www.statsoftinc.com/textbook/stevnet.html.
Stephen Jones, "Neural Networks and the Computation Brain or Maters relating to Artificial Intelligence," The Brain Project, http://www.culture.com.au/brain_proj/neur_net.htm.
David W. Clark, "An Introduction to Neural Networks"; http://members.home.net/neuralnet/introtonn/index.htm.
"A Basic Introduction to Neural Networks"; http://blizzard.gis.uiuc.edu/htmldocs/Neural/neural.html.
Meyer et al., "Computational neural networks: a general purpose tool for nanotechnology," Abstract, 5th Foresight Conference on Molecular Nanotechnology; http://www.foresight.org/Conferences/MNT05/Abstracts/Meyeabst.html.
Saito et al., "A 1M Synapse Self-Learning Digital Neural Network Chip," ISSCC, pp. 6.5-1 to 6.5-10, IEEE 1998.
Espejo, et al., "A 16×16 Cellular Neural Network Chip for Connected Component Detection," Jun. 30, 1999; http://www.imse.cnm.csic.es/Chipcat/espejo/chip-2.pdf.
Pati et al., "Neural Networks for Tactile Perception," Systems Research Center and Dept. of Electrical Engineering, University of Maryland and U.S Naval Research Laboratory. 1987; http://www.isr.umd.edu/TechReports/ISR/1987/TR_87-123/TR_87-123.phtml.
Osamu Fujita, "Statistical estimation of the number of hidden units for feedforward neural networks," Neural Networks 11 (1998), pp. 851-859.
Abraham Harte, "Liquid Crystals Allow Large-Scale Alignment of Carbon Nanotubes," CURJ (Caltech Undergraduate Research Journal), Nov. 2001, vol. 1, No. 2, pp. 44-49.
"Quantum-Dot Arrays for Computation," ORNL Review vol. 34, No. 2, 2001, pp. 1-5 http://www.ornlgov/ORNLReview/v34_2_01/arrays.htm.
Jabri, M.A. et al., "Adaptive Analog VLSI Neural Systems," Chapman & Hall, London SE1 8HN, UK, 1996, pp. 92-95.
Lipson et al., "Automatic Design and Manufacture of Robotic Lifeforms," NATURE, vol. 406, Aug. 31, 2000, pp. 974-978.
Kunitoshi Yamamoto, et al., "Rapid Communication Orientation and Purification of Carbon Nanotubes Using AC Electrophoresis", J. Phys. D. Appl. Phys 31 (1998) L34-L36.
E.S. Snow, et al., "Random networks of carbon nanotubes as electronic material", Applied Physics Letters, vol. 82, No. 12, Mar. 31, 2003, pp. 2145-2147.
R. Martel, et al., "Ambipolar Electrical Transport in Semiconducting Single-Wall Carbon Nanotubes," Physical Review Letters, vol. 87, No. 25, Dec. 17, 2001, pp. 256805-1 to 256805-4.
S. Heinze, et al., "Carbon Nanotubes as Schottky Barrier Transistors", vol. 89, No. 10, Sep. 2, 2002, pp. 106801-1 to 106801-4.
M. Dubson, et al., "Measurement of the conductivity exponent in two-dimensional percolating networks: square lattice versus random-void continuum", Physical Review B, vol. 32, No. 11, Dec. 1, 1985, pp. 7621-7623.
D.J. Frank, et al., "Highly efficient algorithm for percolative transport studies in two dimensions", Physical Review B, vol. 37, No. 1, Jan. 1, 1988, pp. 302-307.
Uma R. Karmarkar, et al., "Mechanisms and significance of spike-timing dependent plasticity," Biol. Cybern. 87, 373-382 (2002), Jan. 28, 2002.
Uma R. Karmarkar, et al., "A Model of Spike-Timing Dependent Plasticity: One or Two Coincidence Detectors?", J. Neurophysiol, vol. 88, pp. 507-513, Jul. 2002.
M.C.W. van Rossum, et al., "Stable Hebbian Learning from Spkke-Timing-Dependent Plasticity", The Journal of Neuroscience, Dec. 1, 2003, 20(23), pp. 8812-8821.

Xiaohui Xie, et al., "Spike-based learning rules and stabilization of persistent neural activity,".
Nace L. Golding, et al., "Dendritic spikes as a mechanism for cooperative long-term potentiation", NATURE, vol. 418, Jul. 18, 2002, pp. 326-330.
Ozgur Turel, et al., "Possible nanoelectronic implementation of neuromorphic networks", Dept. o f Physics and Astronomy, Stony Brook University.
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3; Sep. 9, 2003; American Chemical Society, pp. 1379-1382.
J.M. Tour, et al., "NanoCell Electronic Memories," J.Am. Chem.Soc. 2003, 125, pp. 13279-13283.
J. Zaumseil, et al., "Three-Dimensional and Multilayer Nanostructures Formed by Nanotransfer Printing," Nano Letters, 2003, vol. 3, No. 9; Jul. 31, 2003, American Chemical Society, pp. 1223-1227.
Charles D. Schaper, "Patterned Transfer of Metallic Thin Film Nanostructures by Water-Soluble Polymer Templates," Nano Letters, 2003, vol. 3, No. 9; Jul. 26, 2003, American Chemical Society, pp. 1305-1309.
C.A. Dyke, et al., "Unbundled and Highly Functionalized Carbon Nanotubes from Aqueous Reactions," Nano Letters, 2003, vol. 3, No. 9; Aug. 19, 2003, American Chemical Society, pp. 1215-1218.
J. Chung, et al., "Nanoscale Gap Fabrication by Carbon Nanotube-Extracted Lithography (CEL)", Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003, American Chemical Society, pp. 1029-1031.
O. Harnack, et al., "Rectifiying Behavior of Electrically Aligned ZnO Nanorods," Nano Letters, 2003, vol. 3, No. 8; Jun. 24, 2003, American Chemical Society, pp. 1097-1101.
M. S. Kumar, et al., "Influence of electric field type on the assembly of single walled carbon nanotubes," Chemical Physics Letters 383 (2004), Dec. 2, 2003; pp. 235-239.
S.W. Lee, et al., "Dielectrophoresis and electrohydrodynamics-mediated fluidic assembly of silicon resistors," Applied Physics Letters, vol. 83, No. 18, Nov. 3, 2003, pp. 3833-3835.
R. Krupke, et al., "Simultaneous Deposition of Metallic Bundles of Single-walled Carbon Nanotubes Using Ac-dielectrophoresis," Nano Letters, 2003, vol. 3, No. 8; Jul. 9, 2003; American Chemical Society, pp. 1019-1023.
K. Bradley, et al., "Flexible Nanotube Electronics," Nano Letters, 2003, vol. 3, No. 10; Aug. 9, 2003, America Chemcial Society, pp. 1353-1355.
T.B. Jones, "Frequency-dependent orientation of isolated particle chains," Journal of Electrostatics, 25 (1990), Elsevier Science Publishers, pp. 231-244.
L.A. Nagahara, "Directed placement of suspended carbon nanotubes for nanometer-scale assembly," Applied Physics Letters, vol. 80, No. 20; May 20, 2003; pp. 3826-3628.
A. Bezryadin, et al., "Electrostatic trapping of single conducting nanoparticles between electrodes," Applied Physics Letters, 71 (9), Sep. 1, 1997, pp. 1273-1275.
S. Suzuki, et al., "Quantitative Analysis of DNA Orientation in Stationary AC Electric Fields Using Fluorescence Anisotropy," IEEE Transactions of Industry Applications, vol. 34, No. 1; Jan./Feb. 1998, pp. 75-83.
Phaedon Avouris, "Molecular Electronics with Carbon Nanotubes," Accounts of Chemical Research, vol. 35, No. 12; Jul. 31, 2002, pp. 1025-1034.
X. Liu, et al., "Electric-Field Induced Accumulation and Alignment of Carbon Nanotubes," 2002 Annual Report Conference on Electrical Insulation and Dielectric Phenomena, pp. 31-34.
R. Krupke, et al. "Contacting single bundles of carbon nanotubes with alternating electric fields," Appl. Phys. A. 76, Oct. 28, 2002, pp. 397-400.
M. Law, et al., "Photochemical Sensing of NO_2 with SnO^2 Nanoribbon Nanosensors at Room Temperature," Angew. Chem. 2002, 114, Nr. 13, pp. 2511-2514.
J. Tour, et al., "Nanocell Logic Gates for Molecular Computing," IEEE Transactions on Nanotechnology, vol. 1, No. 2, Jun. 2002, pp. 100-109.
A. Leonardi, et al., "Simulation methodology for dielectrophoresis in microelectronic Lab-on-a-chip," Modeling and Simulation of Microsystems 2002, pp. 96-99.
J. Chung, et al., "Nanoscale Gap Fabrication and Integration of Carbon Nanotubes by Micromachining," Solid-State Sensor, Actuator and Microsystems Workshop, Jun. 2-6, 2003; Hilton Head Island, South Carolina, pp. 161-164.
L. Zheng, et al., "Towards Single Molecule Manipulation with Dielectrophoresis Using Nanoelectrodes," IEEE-NANO 2003, Aug. 12-14, 2003, Moscone Convention Center, San Francisco, CA; pp. 437-440, http://ieeenano2003.arc.nasa.gov/program_contents.pdf.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628
V.C. Moore, et al., "Individually Suspended Single-Walled Carbon Nanotubes in Various Surfactants," Nano Letters, 2003, vol. 3, No. 10; American Chemical Society; Sep. 8, 2003; pp. 1379-1382.
R. Krupke, "Separation of Metallic from Semiconducting Single-Walled Carbon Nanotubes," Science, vol. 301; Jul. 18, 2003; pp. 344-347.
Wolfgang Maass, "On the Relevance of Time in Neural Computation and Learning," In M. Li and A. Maruoka, editors, *Proc. of the 8th International Conference on Algorithmic Learning Theory in Sendai* (*Japan*), vol. 1316 of *Lecture Notes in Computer Science*, pp. 364-388. Springer (Berlin), 1997.
Wolfgang Maass, "Noisy Spiking Neurons with Temporal Coding have more Computational Power than Sigmoidal Neurons," In M. Mozer, M. I. Jordan, and T. Petsche, editors, *Advances in Neural Information Processing Systems*, vol. 9, pp. 211-217. MIT Press (Cambridge), 1997. (pp. 1-13, including Appendix).
L. Perrinet, et al., "Emergence of filters from natural scenes in a sparse spike coding scheme," Neurocomputing, 2003, pp. 1-14, http://www.laurent.perrinet.free.fr/publi/perrinet03neurocomputing.pdf.
L. Perrinet, et al., "Coherence detection in a spiking neuron via Hebbian learning," Neurocomputing, 2002, vol. 44-46, No. C., pp. 817-822, http://www.laurent.perrinet.free.fr/publi/perrinet02.pdf.
A. Jarosz, et al., "An Introductory Note on Gaussian Correlated Random Matrix," Feb. 21, 2003, pp. 1-20 http://www.if.uj.edu.pl/pl/koloSMP/prace/rndmatrix.pdf.
K. Bradley, et al., "Influence of Mobile Ions on Nanotube Based FET Devices," Nano Letters, 2003, vol. 3, No. 5; American Chemical Society, Apr. 4, 2003; pp. 639-641.
A. van Schaik, "Building blocks for electronic spiking neural networks," Neural Networks 14 (2001), pp. 617-628.

* cited by examiner

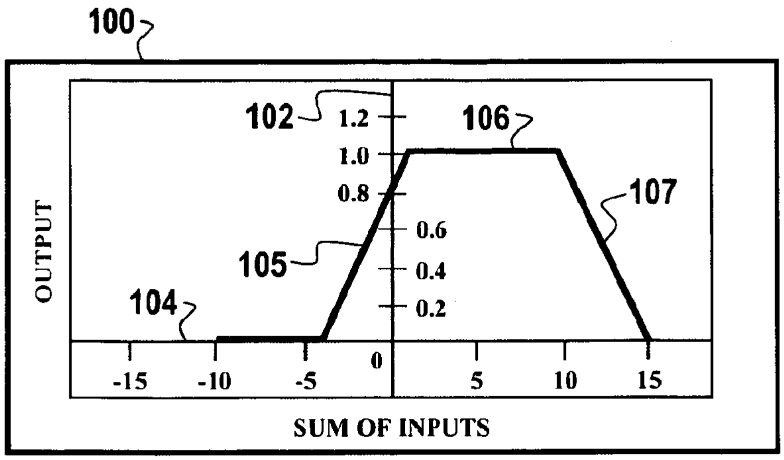


Figure 1

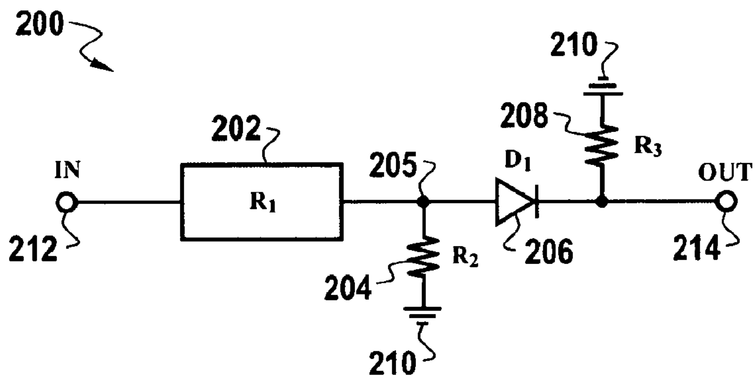


Figure 2

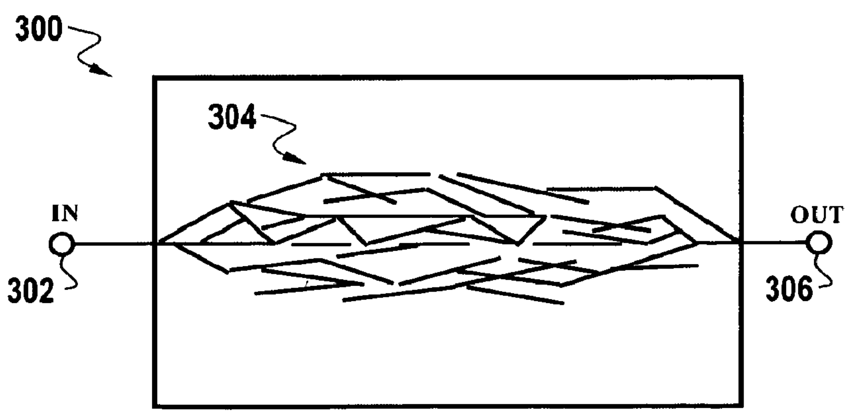


Figure 3

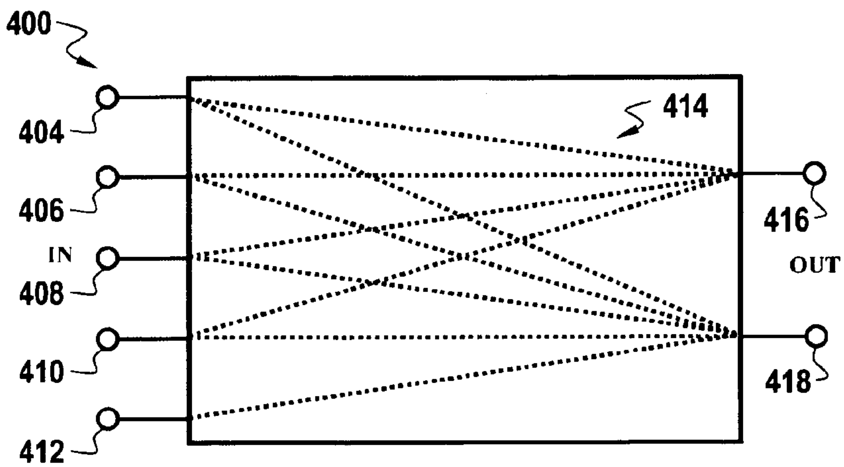


Figure 4

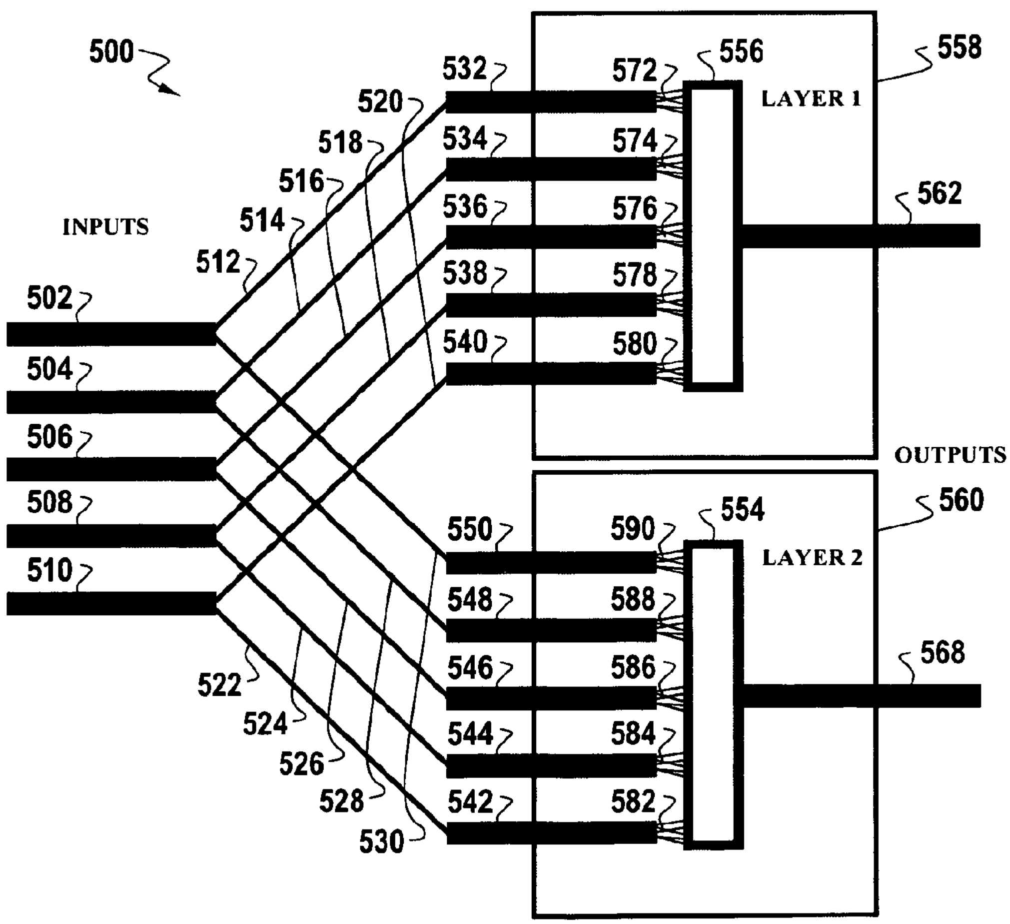


Figure 5

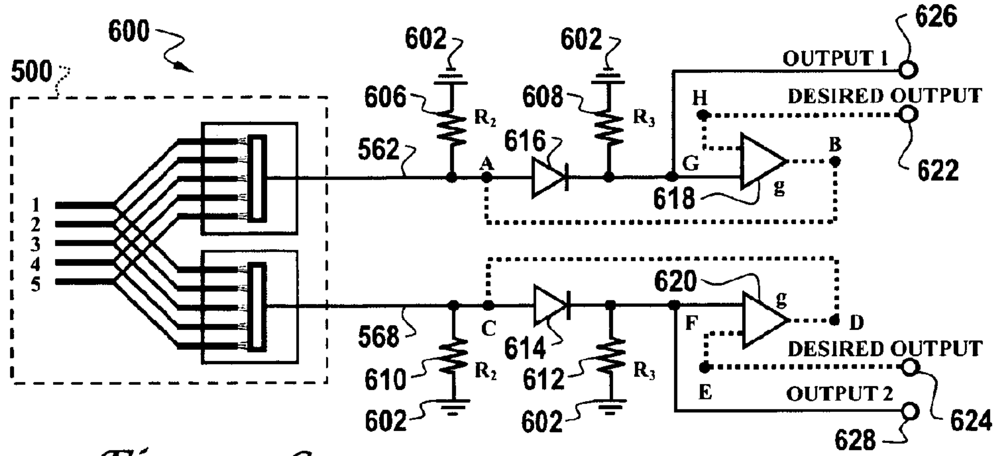
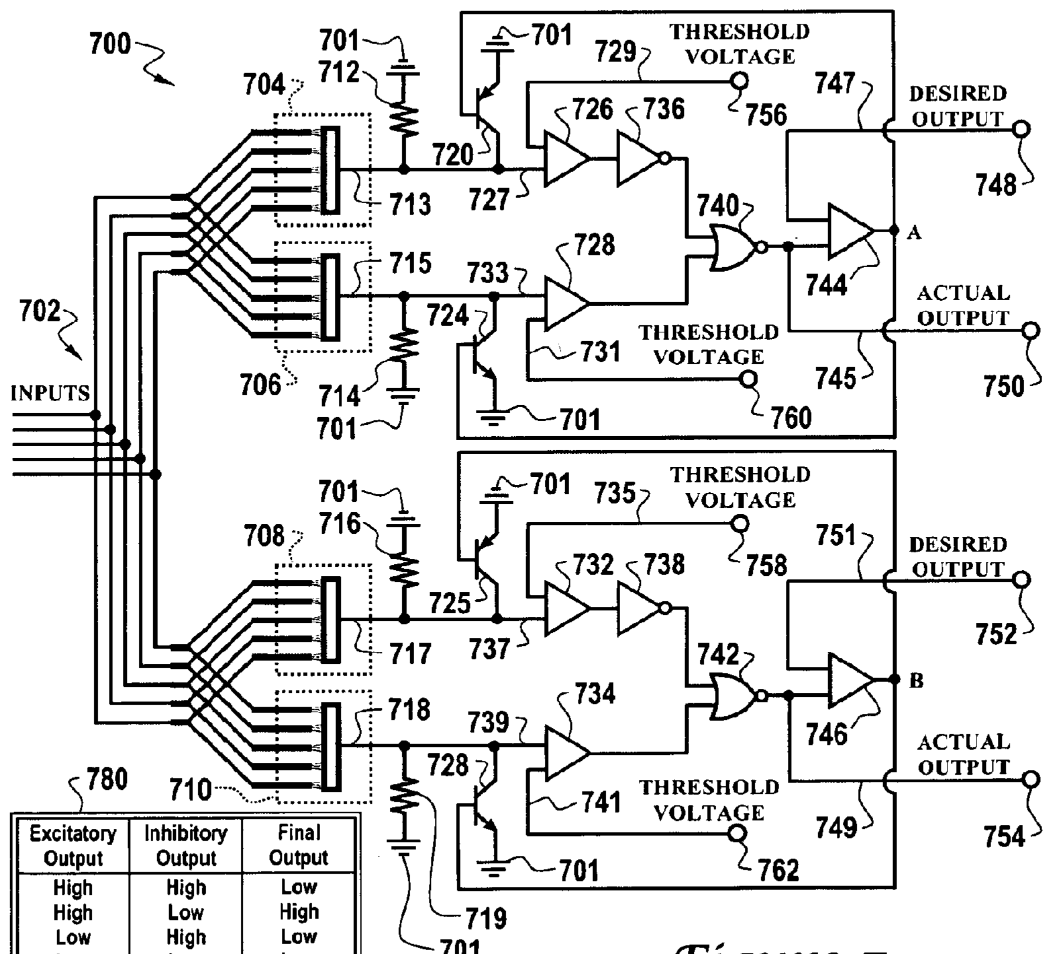


Figure 6



Excitatory Output	Inhibitory Output	Final Output
High	High	Low
High	Low	High
Low	High	Low
Low	Low	Low

Figure 7

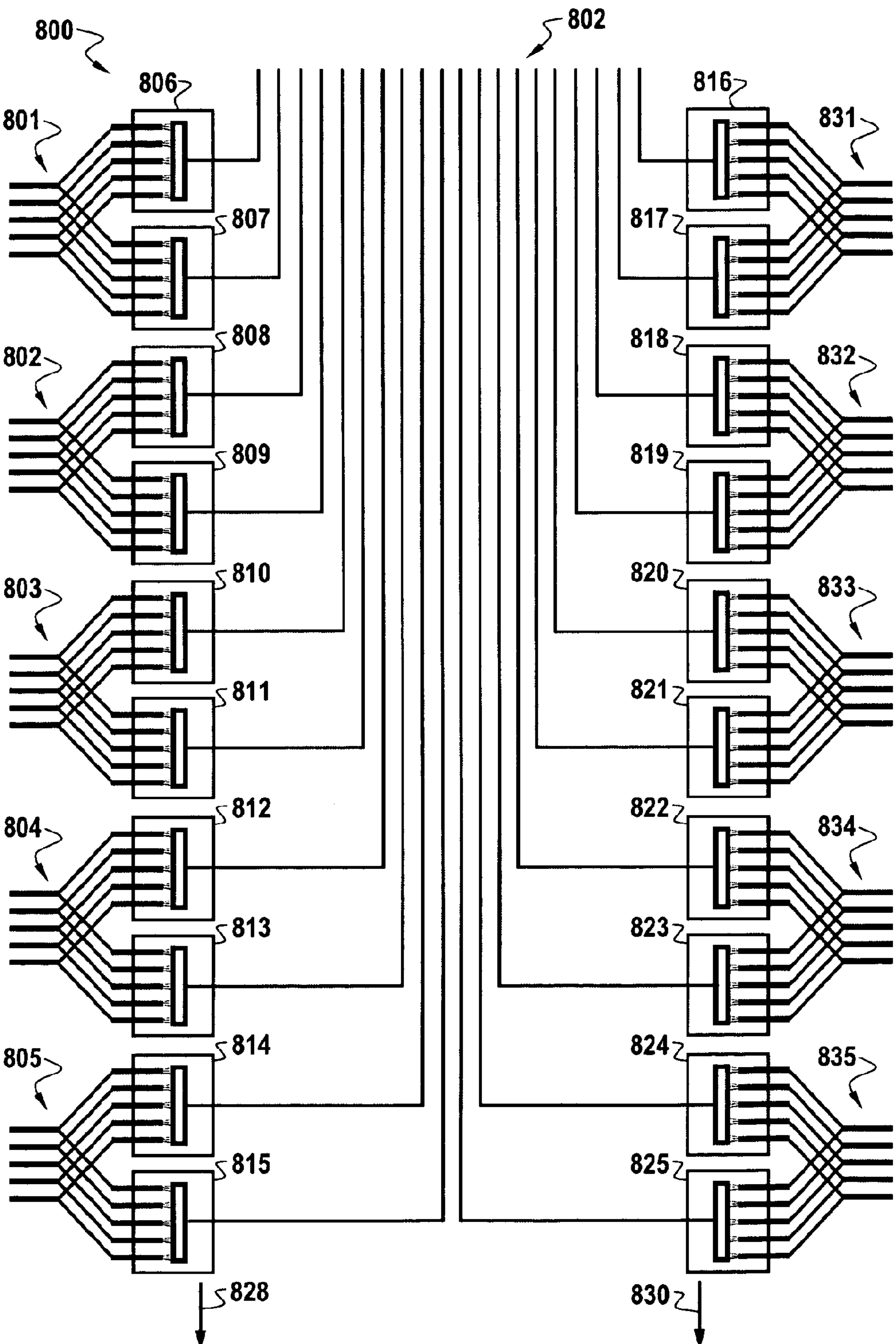

800
802
806
801
807
808
802
809
810
803
811
812
804
813
814
805
815
828
816
831
817
818
832
819
820
833
821
822
834
823
824
835
825
830

Figure 8

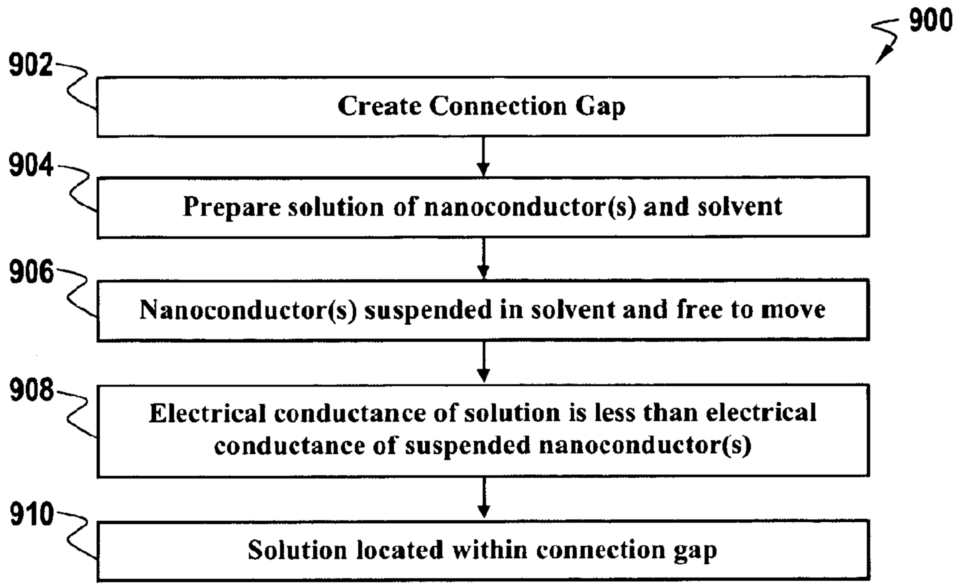


Figure 9

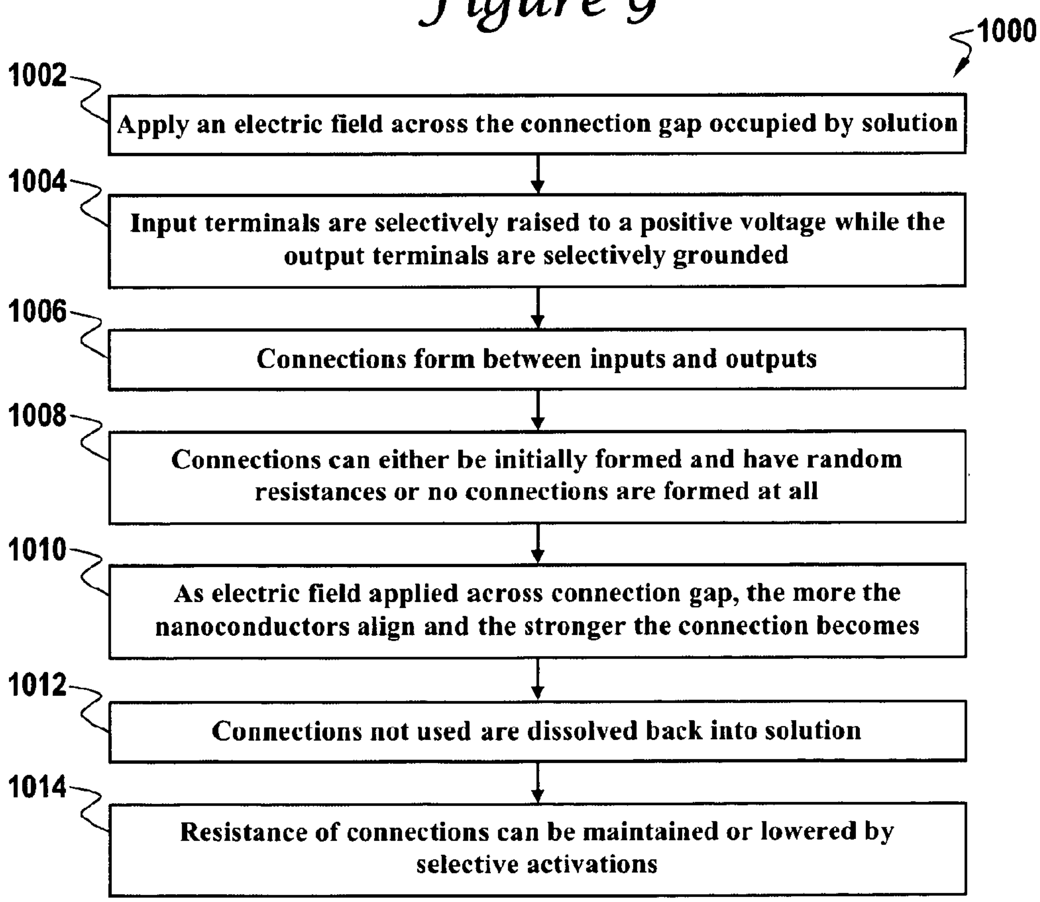


Figure 10

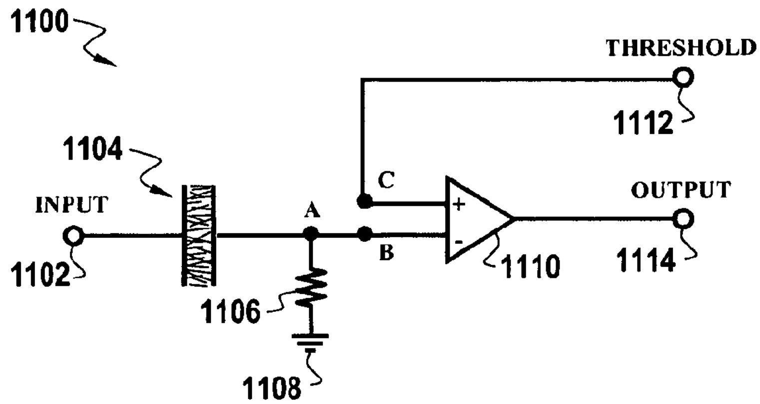


Figure 11

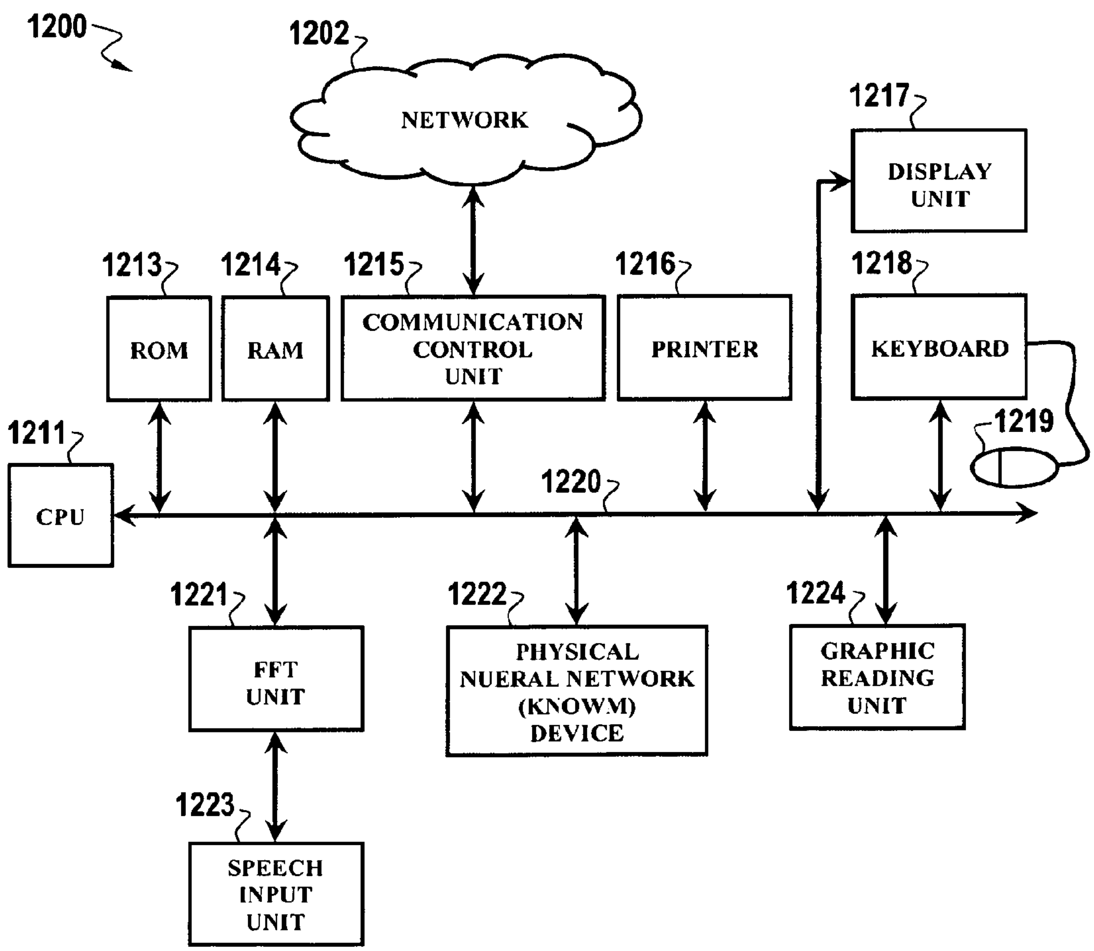


Figure 12

VARIABLE RESISTOR APPARATUS FORMED UTILIZING NANOTECHNOLOGY

CROSS REFERENCE TO RELATED PATENT APPLICATION

This patent application is a continuation of U.S. patent application Ser. No. 10/095,273 entitled "Physical Neural Network Design Incorporating Nanotechnology," which was filed on Mar. 12, 2002 now U.S. Pat. No. 6,889,216, the disclosure of which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

Embodiments generally relate to nanotechnology. Embodiments also relate to variable resistor components. Embodiments additionally relate to neural networks and neural computing systems.

BACKGROUND

Neural networks are computational systems that permit computers to essentially function in a manner analogous to that of the human brain. Neural networks do not utilize the traditional digital model of manipulating 0's and 1's. Instead, neural networks create connections between processing elements, which are equivalent to neurons of a human brain. Neural networks are thus based on various electronic circuits that are modeled on human nerve cells (i.e., neurons). Generally, a neural network is an information-processing network, which is inspired by the manner in which a human brain performs a particular task or function of interest. Computational or artificial neural networks are thus inspired by biological neural systems. The elementary building block of biological neural systems is of course the neuron, the modifiable connections between the neurons, and the topology of the network.

Biologically inspired artificial neural networks have opened up new possibilities to apply computation to areas that were previously thought to be the exclusive domain of human intelligence. Neural networks learn and remember in ways that resemble human processes. Areas that show the greatest promise for neural networks, such as pattern classification tasks such as speech and image recognition, are areas where conventional computers and data-processing systems have had the greatest difficulty.

In general, artificial neural networks are systems composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. The computational elements, or nodes, are connected via variable weights that are typically adapted during use to improve performance. Thus, in solving a problem, neural net models can explore many competing hypothesis simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights. In contrast, with conventional von Neumann computers, an algorithm must first be developed manually, and a program of instructions written and executed sequentially. In some applications, this has proved extremely difficult. This makes conventional computers unsuitable for many real-time problems.

In a neural network, "neuron-like" nodes can output a signal based on the sum of their inputs, the output being the result of an activation function. In a neural network, there exists a plurality of connections, which are electrically coupled among a plurality of neurons. The connections serve as communication bridges among of a plurality of neurons coupled thereto. A network of such neuron-like nodes has the ability to process information in a variety of useful ways. By adjusting the connection values between neurons in a network, one can match certain inputs with desired outputs.

One does not program a neural network. Instead, one "teaches" a neural network by examples. Of course, there are many variations. For instance, some networks do not require examples and extract information directly from the input data. The two variations are thus called supervised and unsupervised learning. Neural networks are currently used in applications such as noise filtering, face and voice recognition and pattern recognition. Neural networks can thus be utilized as an advanced mathematical technique for processing information.

Neural networks that have been developed to date are largely software-based. A true neural network (e.g., the human brain) is massively parallel (and therefore very fast computationally) and very adaptable. For example, half of a human brain can suffer a lesion early in its development and not seriously affect its performance. Software simulations are slow because during the learning phase a standard computer must serially calculate connection strengths. When the networks get larger (and therefore more powerful and useful), the computational time becomes enormous. For example, networks with 10,000 connections can easily overwhelm a computer. In comparison, the human brain has about 100 billion neurons, each of which can be connected to about 5,000 other neurons. On the other hand, if a network is trained to perform a specific task, perhaps taking many days or months to train, the final useful result can be etched onto a piece of silicon and also mass-produced.

A number of software simulations of neural networks have been developed. Because software simulations are performed on conventional sequential computers, however, they do not take advantage of the inherent parallelism of neural network architectures. Consequently, they are relatively slow. One frequently used measurement of the speed of a neural network processor is the number of interconnections it can perform per second. For example, the fastest software simulations available can perform up to about 18 million interconnects per second. Such speeds, however, currently require expensive super computers to achieve. Even so, 18 million interconnects per second is still too slow to perform many classes of pattern classification tasks in real time. These include radar target classifications, sonar target classification, automatic speaker identification, automatic speech recognition and electro-cardiogram analysis, etc.

The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions. Due to the difficulties in building such highly interconnected processors, the currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and built to implement one particular neural network architecture and are not easily, if at all, reconfigurable to implement different architectures. A true physical neural network (i.e., artificial neural network) chip, for example, has not yet been designed and successfully implemented.

The problem with pure hardware implementation of a neural network with technology as it exists today, is the

inability to physically form a great number of connections and neurons. On-chip learning can exist, but the size of the network would be limited by digital processing methods and associated electronic circuitry. One of the difficulties in creating true physical neural networks lies in the highly complex manner in which a physical neural network must be designed and built. It is believed that solutions to creating a true physical and artificial neural network lie in the use of nanotechnology and the implementation of analog variable connections.

The term "Nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

Based on the foregoing, it is believed that a physical neural network which incorporates nanotechnology is a solution to the problems encountered by prior art neural network solutions. Additionally, it is believed that a variable resistor apparatus can be constructed based on nanotechnology and utilized either as an individual component for variable resistance purposes, or in association with physical neural networks, including artificial neurons and components thereof as described herein.

BRIEF SUMMARY

The following summary is provided to facilitate an understanding of some of the innovative features unique to the embodiments, and is not intended to be a full description. A full appreciation of the various aspects of the embodiments can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

It is, therefore, one aspect of the present invention to provide for a variable resistor apparatus that can be formed based on nanotechnology.

It is another aspect of the present invention to provide a physical neural network, which can be formed from a plurality of interconnected nanoconnections or nanoconnectors.

It is a further aspect of the present invention to provide neuron like nodes, which can be formed and implemented utilizing nanotechnology;

It is also an aspect of the present invention to provide a physical neural network that can be formed from one or more neuron-like nodes.

It is yet a further aspect of the present invention to provide a physical neural network, which can be formed from a plurality of nanoconductors, such as, for example, nanowires and/or nanotubes.

The above and other aspects can be achieved as is now described. A physical neural network based on nanotechnology is disclosed herein, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, connected to a plurality of interconnected nanoconnections. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes can be strengthened or weakened according to an application of an electric field. Alignment has also been observed with a magnetic field, but electric fields are generally more practical. Note that the connection network is generally associated with one or more of the neuron-like nodes.

The output signal is generally based on a threshold below which the output signal is not generated and above which the output signal is generated. The transition from zero output to high output need not necessarily be abrupt or non linear. The connection network comprises a number of layers of nanoconnections, wherein the number of layers is generally equal to a number of desired outputs from the connection network. The nanoconnections are formed without influence from disturbances resulting from other nanoconnections thereof. Such nanoconnections may be formed from an electrically conducting material. The electrically conducting material can be selected such that a dipole is induced in the electrically conducting material in the presence of an electric field. Such a nanoconnection may comprise a nanoconductor.

The connection network itself may comprise a connection network structure having a connection gap formed therein, and a solution located within the connection gap, such that the solution comprises a solvent or suspension and one or more nanoconductors. Preferably, a plurality of nanoconductors is present in the solution (i.e., mixture). Note that such a solution may comprise a liquid and/or gas. An electric field can then be applied across the connection gap to permit the alignment of one or more of the nanoconductors within the connection gap. The nanoconductors can be suspended in the solvent, or can lie at the bottom of the connection gap on the surface of the chip. Studies have shown that nanotubes can align both in the suspension and/or on the surface of the gap. The electrical conductance of the mixture is less than the electrical conductance of the nanoconductors within the solution.

The nanoconductors within the connection gap thus experience an increased alignment in accordance with an increase in the electric field applied across the connection gap. Thus, nanoconnections of the neuron-like node that are utilized most frequently by the neuron-like node become stronger with each use thereof. The nanoconnections that are utilized least frequently become increasingly weak and eventually dissolve back into the solution. The nanoconnections may or may not comprise a resistance, which can be raised or lowered by a selective activation of a nanoconnection. They can be configured as nanoconductors such as, for example, a nanotube or nanowire. An example of a nanotube, which may be implemented in accordance with the invention described herein, is a carbon nanotube, nanowire and/or other nanoparticle. Additionally, such nanoconnections may be configured as a negative connection associated with the neuron-like node.

A variable resistor apparatus is also disclosed which includes a plurality of nanoparticles disposed between two terminals, wherein the plurality of nanoparticles provides an electrical resistance. An electric field applied to the plurality of nanoparticles across the two terminals results in an alignment of the nanoparticles over time and a decrease in the electrical resistance thereby providing a variable resistor apparatus. The electric or electrical field can be applied across the two terminals perpendicular to the plurality of nanoconnections. The nanoparticles can comprise nanoconductors, which can be formed as, for example, nanotubes and/or nanowires. The nanoparticles are generally disposed in a solution within a connection gap formed between the

two terminals. The solution can comprise a solvent and/or a suspension of nanoparticles forming a mixture thereof. The solution can also be provided as a liquid, a gel, and or a gas. The solution may also comprise a dielectric.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** illustrates a graph illustrating a typical activation function that can be implemented in accordance with one embodiment;

FIG. **2** illustrates a schematic diagram illustrating a diode configuration as a neuron, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanowires between two electrodes, in accordance with a preferred embodiment;

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections formed between two electrodes, in accordance with a preferred embodiment;

FIG. **4** illustrates a block diagram illustrating a plurality of connections between inputs and outputs of a physical neural network, in accordance with a preferred embodiment;

FIG. **5** illustrates a schematic diagram of a physical neural network that can be created without disturbances, in accordance with a preferred embodiment;

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment;

FIG. **8** illustrates a schematic diagram of a chip layout for a connection network that may be implemented in accordance with an alternative embodiment;

FIG. **9** illustrates a flow chart of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment;

FIG. **10** illustrates a flow chart of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment;

FIG. **11** illustrates a schematic diagram of a circuit illustrating temporal summation within a neuron, in accordance with a preferred embodiment; and

FIG. **12** illustrates a block diagram illustrating a pattern recognition system, which may be implemented with a physical neural network device, in accordance with a preferred embodiment.

DETAILED DESCRIPTION

The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate one or more embodiments.

The physical neural network described and disclosed herein is different from prior art forms of neural networks in that the disclosed physical neural network does not require a computer simulation for training, nor is its architecture based on any current neural network hardware device. The design of the physical neural network described herein with respect to particular embodiments is actually quite "organic". Such a physical neural network is generally fast and adaptable, no matter how large such a physical neural network becomes. The physical neural network described herein can be referred to generically as a Knowm. The terms "physical neural network" and "Knowm" can be utilized interchangeably to refer to the same device, network, or structure.

Network orders of magnitude larger than current VSLI neural networks can be built and trained with a standard computer. One consideration for a Knowm is that it must be large enough for its inherent parallelism to shine through. Because the connection strengths of such a physical neural network are dependant on the physical movement of nanoconnections thereof, the rate at which a small network can learn is generally very small and a comparable network simulation on a standard computer can be very fast. On the other hand, as the size of the network increases, the time to train the device does not change. Thus, even if the network takes a full second to change a connection value a small amount, if it does the same to a billion connections simultaneously, then its parallel nature begins to express itself.

A physical neural network (i.e., a Knowm) must have two components to function properly. First, the physical neural network must have one or more neuron-like nodes that sum a signal and output a signal based on the amount of input signal received. Such a neuron-like node is generally nonlinear in its output. In other words, there should be a certain threshold for input signals, below which nothing is output and above which a constant or nearly constant output is generated or allowed to pass. This is a very basic requirement of standard software-based neural networks, and can be accomplished by an activation function. The second requirement of a physical neural network is the inclusion of a connection network composed of a plurality of interconnected connections (i.e., nanoconnections). Such a connection network is described in greater detail herein.

FIG. **1** illustrates a graph **100** illustrating a typical activation function that can be implemented in accordance with one embodiment. Note that the activation function need not be non-linear, although non-linearity is generally desired for learning complicated input-output relationships. The activation function depicted in FIG. **1** comprises a linear function, and is shown as such for general edification and illustrative purposes only. As explained previously, an activation function may also be non-linear.

As illustrated in FIG. **1**, graph **100** includes a horizontal axis **104** representing a sum of inputs, and a vertical axis **102** representing output values. A graphical line **106** indicates threshold values along a range of inputs from approximately −10 to +10 and a range of output values from approximately 0 to 1. As more neural networks (i.e., active inputs) are established, the overall output as indicated at line **105** climbs until the saturation level indicated by line **106** is attained. If a connection is not utilized, then the level of output (i.e., connection strength) begins to fade until it is revived. This phenomenon is analogous to short term memory loss of a human brain. Note that graph **100** is presented for generally illustrative and edification purposes only and is not considered a limiting feature of the embodiments.

In a Knowm network, the neuron-like node can be configured as a standard diode-based circuit, the diode being the most basic semiconductor electrical component, and the signal it sums may be a voltage. An example of such an arrangement of circuitry is illustrated in FIG. **2**, which generally illustrates a schematic diagram illustrating a diode-based configuration as a neuron **200**, in accordance with a preferred embodiment. Those skilled in the art can appreciate that the use of such a diode-based configuration is not considered a limitation of the embodiments, but merely represents one potential arrangement in which the embodiments may be implemented.

Although a diode may not necessarily be utilized, its current versus voltage characteristics are non-linear when used with associated resistors and similar to the relationship depicted in FIG. 1. The use of a diode as a neuron is thus not a limiting feature, but is only referenced herein with respect to a preferred embodiment. The use of a diode and associated resistors with respect to a preferred embodiment simply represents one potential "neuron" implementation. Such a configuration can be said to comprise an artificial neuron. It is anticipated that other devices and components may be utilized instead of a diode to construct a physical neural network and a neuron-like node (i.e., artificial neuron), as indicated here.

Thus, neuron **200** comprises a neuron-like node that may include a diode **206**, which is labeled D_1, and a resistor **204**, which is labeled R_2. Resistor **204** is connected to a ground **210** and an input **205** of diode **206**. Additionally, a resistor **202**, which is represented as a block and labeled R_1 can be connected to input **205** of diode **206**. Block **202** includes an input **212**, which comprises an input to neuron **200**. A resistor **208**, which is labeled R_3, is also connected to an output **214** of diode **206**. Additionally, resistor **208** is coupled to ground **210**. Diode **206** in a physical neural network is analogous to a neuron of a human brain, while an associated connection formed thereof, as explained in greater detail herein, is analogous to a synapse of a human brain.

As depicted in FIG. **2**, the output **214** is determined by the connection strength of R_1 (i.e., resistor **202**). If the strength of R_1's connection increases (i.e., the resistance decreases), then the output voltage at output **214** also increases. Because diode **206** conducts essentially no current until its threshold voltage (e.g., approximately 0.6V for silicon) is attained, the output voltage will remain at zero until R_1 conducts enough current to raise the pre-diode voltage to approximately 0.6V. After 0.6V has been achieved, the output voltage at output **214** will increase linearly. Simply adding extra diodes in series or utilizing different diode types may increase the threshold voltage.

An amplifier may also be added to the output **214** of diode **206** so that the output voltage immediately saturates at the diode threshold voltage, thus resembling a step function, until a threshold value and a constant value above the threshold is attained. R_3 (i.e., resistor **208**) functions generally as a bias for diode **206** (i.e., D_1) and should generally be about 10 times larger than resistor **204** (i.e., R_2). In the circuit configuration illustrated in FIG. **2**, R_1 can actually be configured as a network of connections composed of many inter-connected conducting nanowires (i.e., see FIG. **3**). As explained previously, such connections are analogous to the synapses of a human brain.

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections **304** formed between two electrodes, in accordance with a preferred embodiment. Nanoconnections **304** (e.g., nanoconductors) depicted in FIG. **3** are generally located between input **302** and output **306**. The network of nanoconnections depicted in FIG. **3** can be implemented as a network of nanoconductors. Examples of nanoconductors include devices such as, for example, nanowires, nanotubes, and nanoparticles.

Nanoconnections **304**, which are analogous to the synapses of a human brain, are preferably composed of electrical conducting material (i.e., nanoconductors). It should be appreciated by those skilled in the art that such nanoconductors can be provided in a variety of shapes and sizes without departing from the teachings herein. For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles.

For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles. For example, each nanoconnection within a connection network may be formed from as a chain of carbon nanoparticles. In "Self-assembled chains of graphitized carbon nanoparticles" by Bezryadin et al., Applied Physics Letters, Vol. 74, No. 18, pp. 2699–2701, May 3, 1999, for example, a technique is reported, which permits the self-assembly of conducting nanoparticles into long continuous chains. Thus, nanoconductors which are utilized to form a physical neural network (i.e., Knowm) could be formed from such nanoparticles. It can be appreciated that the Bezryadin et al is referred to herein for general edification and illustrative purposes only and is not considered to limit the embodiments.

It can be appreciated that a connection network as disclosed herein may be composed from a variety of different types of nanoconductors. For example, a connection network may be formed from a plurality of nanoconductors, including nanowires, nanotubes and/or nanoparticles. Note that such nanowires, nanotubes and/or nanoparticles, along with other types of nanoconductors can be formed from materials such as carbon or silicon. For example, carbon nanotubes may comprise a type of nanotube that can be utilized in accordance with one or more embodiments.

As illustrated in FIG. **3**, nanoconnections **304** comprise a plurality of interconnected nanoconnections, which from this point forward, can be referred to generally as a "connection network." An individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire, a nanotube, nanoparticles(s), or any other nanoconducting structures. Nanoconnections **304** may comprise a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections **304** may be formed from a plurality of interconnected nanoparticles. A connection network is thus not one connection between two electrodes, but a plurality of connections between inputs and outputs. Nanotubes, nanowires, nanoparticles and/or other nanoconducting structures may be utilized, of course, to construct nanoconnections **304** between input **302** and input **306**. Although a single input **302** and a single input **306** is depicted in FIG. **3**, it can be appreciated that a plurality of inputs and a plurality of outputs may be implemented in accordance with the embodiments, rather than simply a single input **302** or a single output **306**.

FIG. **4** illustrates a block diagram illustrating a plurality of nanoconnections **414** between inputs **404, 406, 408, 410, 412** and outputs **416** and **418** of a physical neural network, in accordance with a preferred embodiment. Inputs **404, 406, 408, 410**, and **412** can provide input signals to connections **414**. Output signals can then be generated from connections **414** via outputs **416** and **418**. A connection network can therefore be configured from the plurality of connections **414**. Such a connection network is generally associated with one or more neuron-like nodes.

The connection network also comprises a plurality of interconnected nanoconnections, wherein each nanoconnection thereof is strengthened or weakened according to an application of an electric field. A connection network is not possible if built in one layer because the presence of one connection can alter the electric field so that other connections between adjacent electrodes could not be formed. Instead, such a connection network can be built in layers, so

that each connection thereof can be formed without being influenced by field disturbances resulting from other connections. This can be seen in FIG. 5.

FIG. 5 illustrates a schematic diagram of a physical neural network **500** that can be created without disturbances, in accordance with a preferred embodiment. Physical neural network **500** is composed of a first layer **558** and a second layer **560**. A plurality of inputs **502, 504, 506, 508**, and **510** are respectively provided to layers **558** and **560** respectively via a plurality of input lines **512, 514, 516, 518**, and **520** and a plurality of input lines **522, 524, 526, 528**, and **530**. Input lines **512, 514, 516, 518**, and **520** are further coupled to input lines **532, 534, 536, 538**, and **540** such that each line **532, 534, 536, 538**, and **540** is respectively coupled to nanoconnections **572, 574, 576, 578**, and **580**. Thus, input line **532** is connected to nanconnections **572**. Input line **534** is connected to nanoconnections **574**, and input line **536** is connected to nanoconnections **576**. Similarly, input line **538** is connected to nanconnections **578**, and input line **540** is connected to nanoconnections **580**.

Nanconnections **572, 574, 576, 578**, and **580** may comprise nanoconductors such as, for example, nanotubes and/or nanowires. Nanoconnections **572, 574, 576, 578**, and **580** thus comprise one or more nanoconductors. Additionally, input lines **522, 524, 526, 528**, and **530** are respectively coupled to a plurality of input lines **542, 544, 546, 548** and **550**, which are in turn each respectively coupled to nanoconnections **582, 584, 586, 588**, and **590**. Thus, for example, input line **542** is connected to nanoconnections **582**, while input line **544** is connected to nanoconnections **584**. Similarly, input line **546** is connected to nanoconnections **586** and input line **548** is connected to nanoconnections **588**. Additionally, input line **550** is connected to nanconnections **590**. Box **556** and **554** generally represent simply the output and are thus illustrated connected to outputs **562** and **568**. In other words, outputs **556** and **554** respectively comprise outputs **562** and **568**. The aforementioned input lines and associated components thereof actually comprise physical electronic components, including conducting input and output lines and physical nanoconnections, such as nanotubes and/or nanowires.

Thus, the number of layers **558** and **560** equals the number of desired outputs **562** and **568** from physical neural network **500**. In the previous two figures, every input was potentially connected to every output, but many other configurations are possible. The connection network can be made of any electrically conducting material, although the physics of it requires that they be very small so that they will align with a practical voltage. Carbon nanotubes or any conductive nanowire can be implemented in accordance with the physical neural network described herein. Such components can form connections between electrodes by the presence of an electric field. For example, the orientation and purification of carbon nanotubes has been demonstrated using ac electrophoresis in isopropyl alcohol, as indicated in "Orientation and purification of carbon nanotubes using ac electrophoresis" by Yamamoto et al., J. Phys. D: Applied Physics, 31 (1998), 34–36. Additionally, an electric-field assisted assembly technique used to position individual nanowires suspended in an electric medium between two electrodes defined lithographically on an SiO_2 substrate is indicated in "Electric-field assisted assembly and alignment of metallic nanowires," by Smith et al., Applied Physics Letters, Vol. 77, Num. 9, Aug. 28, 2000. Such references are referred to herein for edification and illustrative purposes only.

The only general requirements for the conducting material utilized to configure the nanoconductors are that such conducting material should preferably conduct electricity, and a dipole should preferably be induced in the material when in the presence of an electric field. Alternatively, the nanoconductors utilized in association with the physical neural network described herein can be configured to include a permanent dipole that is produced by a chemical means, rather than a dipole that is induced by an electric field.

Therefore, it should be appreciated by those skilled in the art that a connection network could also be comprised of other conductive particles that may be developed or found useful in the nanotechnology arts. For example, carbon particles (or "dust") may also be used as nanoconductors in place of nanowires or nanotubes. Such particles may include bearings or granule-like particles.

A connection network can be constructed as follows: A voltage is applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture could be made of many things. The only requirements are that the conducting wires must be suspended in the solvent, either dissolved or in some sort of suspension, free to move around; the electrical conductance of the substance must be less than the electrical conductance of the suspended conducting wire; and the viscosity of the substance should not be too much so that the conducting wire cannot move when an electric field is applied.

The goal for such a connection network is to develop a network of connections of just the right values so as to satisfy the particular signal-processing requirement—exactly what a neural network does. Such a connection network can be constructed by applying a voltage across a space occupied by the mixture mentioned. To create the connection network, the input terminals are selectively raised to a positive voltage while the output terminals are selectively grounded. Thus, connections can gradually form between the inputs and outputs. The important requirement that makes the physical neural network functional as a neural network is that the longer this electric field is applied across a connection gap, or the greater the frequency or amplitude, the more nanotubes and/or nanowires and/or particles align and the stronger the connection thereof becomes. Thus, the connections that are utilized most frequently by the physical neural network become the strongest.

The connections can either be initially formed and have random resistances or no connections may be formed at all. By initially forming random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up from scratch. Depending on the rate of connection decay, having initial random connections could prove faster, although not necessarily. The connection network can adapt itself to the requirements of a given situation regardless of the initial state of the connections. Either initial condition will work, as connections that are not used will "dissolve" back into solution. The resistance of the connection can be maintained or lowered by selective activations of the connection. In other words, if the connection is not used, it will fade away, analogous to the connections between neurons in a human brain. The temperature of the solution can also be maintained at a particular value so that the rate that connections fade away can be controlled. Additionally an electric field can be applied perpendicular to the connections to weaken them, or even erase them out altogether (i.e., as in clear, zero, or reformatting of a "disk").

The nanoconnections may or may not be arranged in an orderly array pattern. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, due to the non-linear nature of neural networks, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

FIG. **6** illustrates a schematic diagram illustrating an example of a physical neural network **600** that can be implemented in accordance with an alternative embodiment. Note that in FIGS. **5** and **6**, like parts are indicated by like reference numerals. Thus, physical neural network **600** can be configured, based on physical neural network **500** illustrated in FIG. **5**. In FIG. **6**, inputs **1**, **2**, **3**, **4**, and **5** are indicated, which are respectively analogous to inputs **502**, **504**, **506**, **508**, and **510** illustrated in FIG. **5**. Outputs **562** and **568** are provided to a plurality of electrical components to create a first output **626** (i.e., Output **1**) and a second output **628** (i.e., Output **2**). Output **562** is tied to a resistor **606**, which is labeled R2 and a diode **616** at node A. Output **568** is tied to a resistor **610**, which is also labeled R2 and a diode **614** at node C. Resistors **606** and **610** are each tied to a ground **602**.

Diode **616** is further coupled to a resistor **608**, which is labeled R3, and first output **626**. Additionally, resistor **608** is coupled to ground **602** and an input to an amplifier **618**. An output from amplifier **618**, as indicated at node B and dashed lines thereof, can be tied back to node A. A desired output **622** from amplifier **618** is coupled to amplifier **618** at node H. Diode **614** is coupled to a resistor **612** at node F. Note that resistor **612** is labeled R3. Node F is in turn coupled to an input of amplifier **620** and to second output **628** (i.e., Output **2**). Diode **614** is also connected to second output **628** and an input to amplifier **620** at second output **628**. Note that second output **628** is connected to the input to amplifier **620** at node F. An output from amplifier **620** is further coupled to node D, which in turn is connected to node C. A desired output **624**, which is indicated by a dashed line in FIG. **6**, is also coupled to an input of amplifier **620** at node E.

In FIG. **6**, the training of physical neural network **600** can be accomplished utilizing, for example, op-amp devices (e.g., amplifiers **618** and **620**). By comparing an output (e.g., first output **626**) of physical neural network **600** with a desired output (e.g., desired output **622**), the amplifier (e.g., amplifier **618**) can provide feedback and selectively strengthen connections thereof. For instance, suppose it is desired to output a voltage of +V at first output **626** (i.e., Output **1**) when inputs **1** and **4** are high. When inputs **1** and **4** are taken high, also assume that first output **626** is zero. Amplifier **618** can then compare the desired output (+V) with the actual output (0) and output −V. In this case, −V is equivalent to ground.

The op-amp outputs and grounds the pre-diode junction (i.e., see node A) and causes a greater electric field across inputs **1** and **4** and the layer **1** output. This increased electric field (larger voltage drop) can cause the nanoconductors in the solution between the electrode junctions to align themselves, aggregate, and form a stronger connection between the **1** and **4** electrodes. Feedback can continue to be applied until output of physical neural network **600** matches the desired output. The same procedure can be applied to every output.

In accordance with the aforementioned example, assume that Output **1** was higher than the desired output (i.e., desired output **622**). If this were the case, the op-amp output can be +V and the connection between inputs **1** and **4** and layer one output can be raised to +V. Columbic repulsions between the nanoconductors can force the connection apart, thereby weakening the connection. The feedback will then continue until the desired output is obtained. This is just one training mechanism. One can see that the training mechanism does not require any computations, because it is a simple feedback mechanism.

Such a training mechanism, however, may be implemented in many different forms. Basically, the connections in a connection network must be able to change in accordance with the feedback provided. In other words, the very general notion of connections being strengthened or connections being weakened in a physical system is the essence of a physical neural network (i.e., Knowm). Thus, it can be appreciated that the training of such a physical neural network may not require a "CPU" to calculate connection values thereof. The Knowm can adapt itself. Complicated neural network solutions could be implemented very rapidly "on the fly", much like a human brain adapts as it performs.

The physical neural network disclosed herein thus has a number of broad applications. The core concept of a Knowm, however, is basic. The very basic idea that the connection values between electrode junctions by nanoconductors can be used in a neural network devise is all that required to develop an enormous number of possible configurations and applications thereof.

Another important feature of a physical neural network is the ability to form negative connections. This is an important feature that makes possible inhibitory effects useful in data processing. The basic idea is that the presence of one input can inhibit the effect of another input. In artificial neural networks as they currently exist, this is accomplished by multiplying the input by a negative connection value. Unfortunately, with a Knowm-based device, the connection may only take on zero or positive values under such a scenario.

In other words, either there can be a connection or no connection. A connection can simulate a negative connection by dedicating a particular connection to be negative, but one connection cannot begin positive and through a learning process change to a negative connection. In general, if starts positive, it can only go to zero. In essence, it is the idea of possessing a negative connection initially that results in the simulation, because this does not occur in a human brain. Only one type of signal travels through axon/dendrites in a human brain. That signal is transferred into the flow of a neurotransmitter whose effect on the postsynaptic neuron can be either excitatory or inhibitory, depending on the neuron.

One method for solving this problem is to utilize two sets of connections for the same output, having one set represent the positive connections and the other set represent the negative connections. The output of these two layers can be compared, and the layer with the greater output will output either a high signal or a low signal, depending on the type of connection set (inhibitory or excitatory). This can be seen in FIG. **7**.

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network **700** that can be implemented in accordance with an alternative embodiment. Physical neural network **700** thus comprises a plurality of

inputs **702** (not necessarily binary) which are respectively fed to layers **704, 706, 708**, and **710**. Each layer is analogous to the layers depicted earlier, such as for example layers **558** and **560** of FIG. **5**. An output **713** of layer **704** can be connected to a resistor **712**, a transistor **720** and a first input **727** of amplifier **726**. Transistor **720** is generally coupled between ground **701** and first input **727** of amplifier **726**. Resistor **712** is connected to a ground **701**. Note that ground **701** is analogous to ground **602** illustrated in FIG. **6** and ground **210** depicted in FIG. **2**. A second input **729** of amplifier **726** can be connected to a threshold voltage **756**. The output of amplifier **726** can in turn be fed to an inverting amplifier **736**.

The output of inverting amplifier **736** can then be input to a NOR device **740**. Similarly, an output **716** of layer **706** may be connected to resistor **714**, transistor **733** and a first input **733** of an amplifier **728**. A threshold voltage **760** is connected to a second input **737** of amplifier **728**. Resistor **714** is generally coupled between ground **701** and first input **733** of amplifier **728**. Note that first input **733** of amplifier **728** is also generally connected to an output **715** of layer **706**. The output of amplifier **728** can in turn be provided to NOR device **740**. The output from NOR device **740** is generally connected to a first input **745** of an amplifier **744**. An actual output **750** can be taken from first input **745** to amplifier **744**. A desired output **748** can be taken from a second input **747** to amplifier **744**. The output from amplifier **744** is generally provided at node A, which in turn is connected to the input to transistor **720** and the input to transistor **724**. Note that transistor **724** is generally coupled between ground **701** and first input **733** of amplifier **728**. The second input **731** of amplifier **728** can produce a threshold voltage **760**.

Layer **708** provides an output **717** that can be connected to resistor **716**, transistor **725** and a first input **737** to an amplifier **732**. Resistor **716** is generally coupled between ground **701** and the output **717** of layer **708**. The first input **737** of amplifier **732** is also electrically connected to the output **717** of layer **708**. A second input **735** to amplifier **732** may be tied to a threshold voltage **758**. The output from amplifier **732** can in turn be fed to an inverting amplifier **738**. The output from inverting amplifier **738** may in turn be provided to a NOR device **742**. Similarly, an output **718** from layer **710** can be connected to a resistor **719**, a transistor **728** and a first input **739** of an amplifier **734**. Note that resistor **719** is generally coupled between node **701** and the output **719** of layer **710**. A second input **741** of amplifier **734** may be coupled to a threshold voltage **762**. The output from of NOR device **742** is generally connected to a first input **749** of an amplifier **746**. A desired output **752** can be taken from a second input **751** of amplifier **746**. An actual output **754** can be taken from first input **749** of amplifier **746**. The output of amplifier **746** may be provided at node B, which in turn can be tied back to the respective inputs to transistors **725** and **728**. Note that transistor **725** is generally coupled between ground **701** and the first input **737** of amplifier **732**. Similarly, transistor **728** is generally connected between ground **701** and the first input **739** of amplifier **734**.

Note that transistors **720, 724, 725** and/or **728** each can essentially function as a switch to ground. A transistor such as, for example, transistor **720, 724, 725** and/or **728** may comprise a field-effect transistor (FET) or another type of transistor, such as, for example, a single-electron transistor (SET). Single-electron transistor (SET) circuits are essential for hybrid circuits combining quantum SET devices with conventional electronic devices. Thus, SET devices and circuits may be adapted for use with the physical neural network of the embodiments. This is particularly important because as circuit design rules begin to move into regions of the sub-100 nanometer scale, where circuit paths are only 0.001 of the thickness of a human hair, prior art device technologies will begin to fail, and current leakage in traditional transistors will become a problem. SET offers a solution at the quantum level, through the precise control of a small number of individual electrons. Transistors such as transistors **720, 724, 725** and/or **728** can also be implemented as carbon nanotube transistors.

A truth table for the output of circuit **700** is illustrated at block **780** in FIG. **7**. As indicated at block **780**, when an excitatory output is high and the inhibitory output is also high, the final output is low. When the excitatory output is high and the inhibitory output is low, the final output is high. Similarly, when the excitatory output is low and the inhibitory output is high, the final output is low. When the excitatory output is low and the inhibitory output is also low, the final output is low. Note that layers **704** and **708** may thus comprise excitatory connections, while layers **706** and **710** may comprise inhibitory connections.

For every desired output, two sets of connections are used. The output of a two-diode neuron can be fed into an op-amp (e.g., a comparator). If the output that the op-amp receives is low when it should be high, the op-amp outputs a low signal. This low signal can cause the transistors (e.g., transistors **720, 725**) to saturate and ground out the pre-diode junction for the excitatory diode. Such a scenario can cause, as indicated previously, an increase in the voltage drop across those connections that need to increase their strength. Note that only those connections going to the excitatory diode are strengthened. Likewise, if the desired output were low when the actual output was high, the op-amp can output a high signal. This can cause the inhibitory transistor (e.g., an NPN transistor) to saturate and ground out the neuron junction of the inhibitory connections. Those connections going to the inhibitory diode can thereafter strengthen.

At all times during the learning process, a weak alternating electric field can be applied perpendicular to the connections. This can cause the connections to weaken by rotating the nanotube perpendicular to the connection direction. This perpendicular field is important because it can allow for a much higher degree of adaptation. To understand this, one must realize that the connections cannot (practically) keep getting stronger and stronger. By weakening those connections not contributing much to the desired output, we decrease the necessary strength of the needed connections and allow for more flexibility in continuous training. This perpendicular alternating voltage can be realized by the addition of two electrodes on the outer extremity of the connection set, such as plates sandwiching the connections (i.e., above and below). Other mechanisms, such as increasing the temperature of the nanotube suspension could also be used for such a purpose, although this method is perhaps a little less controllable or practical.

The circuit depicted in FIG. **7** can be separated into two separate circuits. The first part of the circuit can be composed of nanotube connections, while the second part of the circuit comprises the "neurons" and the learning mechanism (i.e., op-amps/comparator). The learning mechanism on first glance appears similar to a relatively standard circuit that could be implemented on silicon with current technology. Such a silicon implementation can thus comprise the "neuron" chip. The second part of the circuit (i.e., the connections) is thus a new type of chip, although it could be constructed with current technology. The connection chip

can be composed of an orderly array of electrodes spaced anywhere from, for example, 100 nm to 1 μm or perhaps even further. In a biological system, one talks of synapses connecting neurons. It is in the synapses where the information is processed, (i.e., the "connection weights"). Similarly, such a chip can contain all of the synapses for the physical neural network. A possible arrangement thereof can be seen in FIG. **8**.

FIG. **8** illustrates a schematic diagram of a chip layout **800** for a connection network that may be implemented in accordance with an alternative embodiment. FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., connection network **800**) that can be implemented in accordance with one or more embodiments. Chip layout **800** includes an input array composed of plurality of inputs **801**, **802**, **803**, **804**, and **805**, which are provided to a plurality of layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. A plurality of outputs **802** can be derived from layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. Inputs **801** can be coupled to layers **806** and **807**, while inputs **802** can be connected to layers **808** and **809**. Similarly, inputs **803** can be connected to layers **810** and **811**. Also, inputs **804** can be connected to layers **812** and **813**. Inputs **805** are generally connected to layers **814** and **815**.

Similarly, such an input array can includes a plurality of inputs **831**, **832**, **833**, **834** and **835** which are respectively input to a plurality of layers **816**, **817**, **818**, **819**, **820**, **821**, **822**, **823**, **824** and **825**. Thus, inputs **831** can be connected to layers **816** and **817**, while inputs **832** are generally coupled to layers **818** and **819**. Additionally, inputs **833** can be connected to layers **820** and **821**. Inputs **834** can be connected to layers **822** and **823**. Finally, inputs **835** are connected to layers **824** and **825**. Arrows **828** and **830** represent a continuation of the aforementioned connection network pattern. Those skilled in the art can appreciate, of course, that chip layout **800** is not intended to represent an exhaustive chip layout or to limit the scope of the invention. Many modifications and variations to chip layout **800** are possible in light of the teachings herein without departing from the scope of the embodiments. It is contemplated that the use of a chip layout, such as chip layout **800**, can involve a variety of components having different characteristics.

Preliminary calculations based on a maximum etching capability of 200 nm resolution indicated that over 4 million synapses could fit on an area of approximately 1 cm^2. The smallest width that an electrode can possess is generally based on current lithography. Such a width may of course change as the lithographic arts advance. This value is actually about 70 nm for state-of-the-art techniques currently. These calculations are of course extremely conservative, and are not considered a limiting feature of the embodiments. Such calculations are based on an electrode with, separation, and gap of approximately 200 nm. For such a calculation, for example, 166 connection networks comprising 250 inputs and 100 outputs can fit within a one square centimeter area.

If such chips are stacked vertically, an untold number of synapses could be attained. This is two to three orders of magnitude greater than some of the most capable neural network chips out there today, chips that rely on standard methods to calculate synapse weights. Of course, the geometry of the chip could take on many different forms, and it is quite possible (based on a conservative lithography and chip layout) that many more synapses could fit in the same space. The training of a chip this size would take a fraction of the time of a comparably sized traditional chip using digital technology.

The training of such a chip is primarily based on two assumptions. First, the inherent parallelism of a physical neural network (i.e., a Knowm) can permit all training sessions to occur simultaneously, no matter now large the associated connection network. Second, recent research has indicated that near perfect aligning of nanotubes can be accomplished in approximately 15 minutes. If one considers that the input data, arranged as a vector of binary "high's" and "low's" is presented to the Knowm simultaneously, and that all training vectors are presented one after the other in rapid succession (e.g., perhaps 100 MHz or more), then each connection would "see" a different frequency in direct proportion to the amount of time that its connection is required for accurate data processing (i.e., provided by a feedback mechanism). Thus, if it only takes approximately 15 minutes to attain an almost perfect state of alignment, then this amount of time would comprise the longest amount of time required to train, assuming that all of the training vectors are presented during that particular time period.

FIG. **9** illustrates a flow chart **900** of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment. Initially, as indicated at block **902**, a connection gap is created from a connection network structures. As indicated earlier, the goal for such a connection network is generally to develop a network of connections of "just" the right values to satisfy particular information processing requirements, which is precisely what a neural network accomplishes. As illustrated at block **904**, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent," and also a suspension.

The solvent utilized can comprise a volatile liquid that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution may be sandwiched between wafers of silicon or other materials. If the fluid has a melting point that is approximately at room temperature, then the viscosity of the fluid could be controlled easily. Thus, if it is desired to lock the connection values into a particular state, the associated physical neural network (i.e., Knowm) may be cooled slightly until the fluid freezes. The term "solvent" as utilized herein thus can include fluids such as for example, toluene, hexadecane, mineral oil, etc. Note that the solution in which the nanoconductors (i.e., nanoconnections) are present should generally comprise a dielectric. Thus, when the resistance between the electrodes is measured, the conductivity of the nanoconductors can be essentially measured, not that of the solvent. The nanoconductors can be suspended in the solution or can alternately lie on the bottom surface of the connection gap. The solvent may also be provided in the form of a gas.

As illustrated thereafter at block **906**, the nanoconductors must be suspended in the solvent, either dissolved or in a suspension of sorts, but generally free to move around, either in the solution or on the bottom surface of the gap. As depicted next at block **908**, the electrical conductance of the solution must be less than the electrical conductance of the suspended nanoconductor(s). Similarly, the electrical resistance of the solution is greater than the electrical resistance of the nanoconductor.

Next, as illustrated at block **910**, the viscosity of the substance should not be too much so that the nanoconductors cannot move when an electric field (e.g., voltage) is applied. Finally, as depicted at block **912**, the resulting

solution of the "solvent" and the nanoconductors is thus located within the connection gap.

Note that although a logical series of steps is illustrated in FIG. **9**, it can be appreciated that the particular flow of steps can be re-arranged. Thus, for example, the creation of the connection gap, as illustrated at block **902**, may occur after the preparation of the solution of the solvent and nanoconductor(s), as indicated at block **904**. FIG. **9** thus represents merely possible series of steps, which may be followed to create a connection network. A variety of other steps may be followed as long as the goal of achieving a connection network is achieved. Similar reasoning also applies to FIG. **10**.

FIG. **10** illustrates a flow chart **1000** of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment. As indicated at block **1002**, an electric field can be applied across the connection gap discussed above with respect to FIG. **9**. The connection gap can be occupied by the solution discussed above. As indicated thereafter at block **1004**, to create the connection network, the input terminals can be selectively raised to a positive voltage while the output terminals are selectively grounded. As illustrated thereafter at block **1006**, connections thus form between the inputs and the outputs. The important requirements that make the resulting physical neural network functional as a neural network is that the longer this electric field is applied across the connection gap, or the greater the frequency or amplitude, the more nanoconductors align and the stronger the connection becomes. Thus, the connections that get utilized the most frequently become the strongest.

As indicated at block **1008**, the connections can either be initially formed and have random resistances or no connections will be formed at all. By forming initial random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up as much. Depending on the rate of connection decay, having initial random connections could prove to be a faster method, although not necessarily. A connection network can adapt itself to whatever is required regardless of the initial state of the connections. Thus, as indicated at block **1010**, as the electric field is applied across the connection gap, the more the nonconductor(s) will align and the stronger the connection becomes. Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block **1012**. As illustrated at block **1014**, the resistance of the connection can be maintained or lowered by selective activations of the connections. In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain.

The neurons in a human brain, although seemingly simple when viewed individually, interact in a complicated network that computes with both space and time. The most basic picture of a neuron, which is usually implemented in technology, is a summing device that adds up a signal. Actually, this statement can be made even more general by stating that a neuron adds up a signal in discrete units of time. In other words, every group of signals incident upon the neuron can be viewed as occurring in one moment in time. Summation thus occurs in a spatial manner. The only difference between one signal and another signal depends on where such signals originate. Unfortunately, this type of data processing excludes a large range of dynamic, varying situations that cannot necessarily be broken up into discrete units of time.

The example of speech recognition is a case in point. Speech occurs in the time domain. A word is understood as the temporal pronunciation of various syllables. A sentence is composed of the temporal separation of varying words. Thoughts are composed of the temporal separation of varying sentences. Thus, for an individual to understand a spoken language at all, a syllable, word, sentence or thought must exert some type of influence on another syllable, word, sentence or thought. The most natural way that one sentence can exert any influence on another sentence, in the light of neural networks, is by a form of temporal summation. That is, a neuron "remembers" the signals it received in the past.

The human brain accomplishes this feat in an almost trivial manner. When a signal reaches a neuron, the neuron has an influx of ions rush through its membrane. The influx of ions contributes to an overall increase in the electrical potential of the neuron. Activation is achieved when the potential inside the cell reaches a certain threshold. The one caveat is that it takes time for the cell to pump out the ions, something that it does at a more or less constant rate. So, if another signal arrives before the neuron has time to pump out all of the ions, the second signal will add with the remnants of the first signal and achieve a raised potential greater than that which could have occurred with only the second signal. The first signal influences the second signal, which results in temporal summation.

Implementing this in a technological manner has proved difficult in the past. Any simulation would have to include a "memory" for the neuron. In a digital representation, this requires data to be stored for every neuron, and this memory would have to be accessed continually. In a computer simulation, one must discritize the incoming data, since operations (such as summations and learning) occur serially. That is, a computer can only do one thing at a time. Transformations of a signal from the time domain into the spatial domain require that time be broken up into discrete lengths, something that is not necessarily possible with real-time analog signals in which no point exists within a time-varying signal that is uninfluenced by another point.

A physical neural network, however, is generally not digital. A physical neural network is a massively parallel analog device. The fact that actual molecules (e.g., nanoconductors) must move around (in time) makes temporal summation a natural occurrence. This temporal summation is built into the nanoconnections. The easiest way to understand this is to view the multiplicity of nanoconnections as one connection with one input into a neuron-like node (Op-amp, Comparator, etc.). This can be seen in FIG. **11**.

FIG. **11** illustrates a schematic diagram of a circuit **1100** illustrating temporal summation within a neuron, in accordance with a preferred embodiment. As indicated in FIG. **11**, an input **1102** can be provided to nanoconnections **1104**, which in turn can provide a signal, which can be input to an amplifier **1110** (e.g., op amp) at node B. A resistor **1106** can be connected to node A, which in turn is electrically equivalent to node B. Node B can be connected to a negative input of amplifier **1100**. Resistor **1108** can also be connected to a ground **1108**. Amplifier **1110** provides output **1114**. Note that although nanoconnections **1104** is referred to in the plural it can be appreciated that nanoconnections **1104** can comprise a single nanoconnection or a plurality of nanoconnections. For simplicity sake, however, the plural form is used to refer to nanoconnections **1104**.

Input **1102** can be provided by another physical neural network (i.e., Knowm) to cause increased connection strength of nanoconnections **1104** over time. This input would most likely arrive in pulses, but could also be continuous. A constant or pulsed electric field perpendicular to the connections can serve to constantly erode the con-

nections, so that only signals of a desired length or amplitude can cause a connection to form. Once the connection is formed, the voltage divider formed by nanoconnection **1104** and resistor **1106** can cause a voltage at node A in direct proportion to the strength of nanoconnections **1104**. When the voltage at node A reaches a desired threshold, the amplifier (i.e., an op-amp and/or comparator), will output a high voltage (i.e., output **1114**). The key to the temporal summation is that, just like a real neuron, it takes time for the electric field to breakdown the nanoconnections **1104**, so that signals arriving close in time will contribute to the firing of the neuron (i.e., op-amp, comparator, etc.). Temporal summation has thus been achieved. The parameters of the temporal summation could be adjusted by the amplitude and frequency of the input signals and the perpendicular electric field.

FIG. **12** illustrates a block diagram illustrating a pattern recognition system **1200**, which may be implemented with a physical neural network device **1222**, in accordance with an alternative embodiment. Note that pattern recognition system **1200** can be implemented as a speech recognition system. Although pattern recognition system **1200** is depicted herein in the context of speech recognition, a physical neural network device (i.e., a Knowm device) may be implemented with other pattern recognition systems, such as visual and/or imaging recognition systems. FIG. **12** thus does not comprise a limiting feature of the embodiments and is presented for general edification and illustrative purposes only. Those skilled in the art can appreciate that the diagram depicted in FIG. **12** may be modified as new applications and hardware are developed. The development or use of a pattern recognition system such as pattern recognition system **1200** of FIG. **12** by no means limits the scope of the physical neural network (i.e., Knowm) disclosed herein.

FIG. **12** thus illustrates in block diagram fashion, the system structure of a speech recognition device using a neural network according to an alternative embodiment. The pattern recognition system **1200** can be provided with a CPU **1211** for performing the functions of inputting vector rows and instructor signals (vector rows) to an output layer for the learning process of a physical neural network device **1222**, and changing connection weights between respective neuron devices based on the learning process. Pattern recognition system **1200** can be implemented within the context of a data-processing system, such as, for example, a personal computer or personal digital assistant (PDA), both of which are well known in the art.

The CPU **1211** can perform various processing and controlling functions, such as pattern recognition, including but not limited to speech and/or visual recognition based on the output signals from the physical neural network device **1222**. The CPU **1211** is connected to a read-only memory (ROM) **1213**, a random-access memory (RAM) **1214**, a communication control unit **1215**, a printer **1216**, a display unit **1217**, a keyboard **1218**, an FFT (fast Fourier transform) unit **1221**, a physical neural network device **1222** and a graphic reading unit **1224** through a bus line **1220** such as a data bus line. The bus line **1220** may comprise, for example, an ISA, EISA, or PCI bus.

The ROM **1213** is a read-only memory storing various programs or data used by the CPU **1211** for performing processing or controlling the learning process, and speech recognition of the physical neural network device **1222**. The ROM **1213** may store programs for carrying out the learning process according to error back-propagation for the physical neural network device or code rows concerning, for example, 80 kinds of phonemes for performing speech recognition. The code rows concerning the phonemes can be utilized as second instructor signals and for recognizing phonemes from output signals of the neuron device network. Also, the ROM **1213** can store programs of a transformation system for recognizing speech from recognized phonemes and transforming the recognized speech into a writing (i.e., written form) represented by characters.

A predetermined program stored in the ROM **1213** can be downloaded and stored in the RAM **1214**. RAM **1214** generally functions as a random access memory used as a working memory of the CPU **1211**. In the RAM **1214**, a vector row storing area can be provided for temporarily storing a power obtained at each point in time for each frequency of the speech signal analyzed by the FFT unit **1221**. A value of the power for each frequency serves as a vector row input to a first input portion of the physical neural network device **1222**. Further, in the case where characters or graphics are recognized in the physical neural network device, the image data read by the graphic reading unit **1224** are stored in the RAM **1214**.

The communication control unit **1215** transmits and/or receives various data such as recognized speech data to and/or from another communication control unit through a communication network **1202** such as a telephone line network, an ISDN line, a LAN, or a personal computer communication network. Network **1202** may also comprise, for example, a telecommunications network, such as a wireless communications network. Communication hardware methods and systems thereof are well known in the art.

The printer **1216** can be provided with a laser printer, a bubble-type printer, a dot matrix printer, or the like, and prints contents of input data or the recognized speech. The display unit **1217** includes an image display portion such as a CRT display or a liquid crystal display, and a display control portion. The display unit **1217** can display the contents of the input data or the recognized speech as well as a direction of an operation required for speech recognition utilizing a graphical user interface (GUI).

The keyboard **1218** generally functions as an input unit for varying operating parameters or inputting setting conditions of the FFT unit **1221**, or for inputting sentences. The keyboard **1218** is generally provided with a ten-key numeric pad for inputting numerical figures, character keys for inputting characters, and function keys for performing various functions. A mouse **1219** can be connected to the keyboard **1218** and serves as a pointing device.

A speech input unit **1223**, such as a microphone can be connected to the FFT unit **1221**. The FFT unit **1221** transforms analog speech data input from the voice input unit **1223** into digital data and carries out spectral analysis of the digital data by discrete Fourier transformation. By performing a spectral analysis using the FFT unit **1221**, the vector row based on the powers of the respective frequencies are output at predetermined intervals of time. The FFT unit **1221** performs an analysis of time-series vector rows, which represent characteristics of the inputted speech. The vector rows output by the FFT **1221** are stored in the vector row storing area in the RAM **1214**.

The graphic reading unit **224**, provided with devices such as a CCD (Charged Coupled Device), can be used for reading images such as characters or graphics recorded on paper or the like. The image data read by the image-reading unit **1224** are stored in the RAM **1214**. Note that an example of a pattern recognition apparatus, which may be modified for use with the physical neural network described herein, is disclosed in U.S. Pat. No. 6,026,358 to Tomabechi, Feb. 16, 2000, "Neural Network, A Method of Learning of a Neural

Network and Phoneme Recognition Apparatus Utilizing a Neural Network." U.S. Pat. No. 6,026,358 is incorporated herein by reference. It can be appreciated that the Tomabechi reference does not teach, suggest or anticipate the embodiments, but is discussed herein for general illustrative, background and general edification purposes only.

The implications of a physical neural network are tremendous. With existing lithography technology, many electrodes in an array such as depicted in FIG. **5** can be etched onto a wafer of silicon. The neuron-diodes, as well as the training circuitry illustrated in FIG. **6**, could be built onto the same silicon wafer, although it may be desirable to have the connections on a separate chip due to the liquid solution of nanoconductors. A solution of suspended nanoconductors could be placed between the electrode connections and the chip could be packaged. The resulting "chip" would look much like a current Integrated Chip (IC) or VLSI (very large scale integrated) chips. One could also place a rather large network parallel with a computer processor as part of a larger system. Such a network, or group of networks, could add significant computational capabilities to standard computers and associated interfaces.

For example, such a chip may be constructed utilizing a standard computer processor in parallel with a large physical neural network or group of physical neural networks. A program can then be written such that the standard computer teaches the neural network to read, or create an association between words, which is precisely the same sort of task in which neural networks can be implemented. Once the physical neural network is able to read, it can be taught for example to "surf" the Internet and find material of any particular nature. A search engine can then be developed that does not search the Internet by "keywords", but instead by meaning. This idea of an intelligent search engine has already been proposed for standard neural networks, but until now has been impractical because the network required was too big for a standard computer to simulate. The use of a physical neural network (i.e., physical neural network) as disclosed herein now makes a truly intelligent search engine possible.

A physical neural network can be utilized in other applications, such as, for example, speech recognition and synthesis, visual and image identification, management of distributed systems, self-driving cars, filtering, etc. Such applications have to some extent already been accomplished with standard neural networks, but are generally limited in expense, practicality and not very adaptable once implemented. The use of a physical neural network can permit such applications to become more powerful and adaptable. Indeed, anything that requires a bit more "intelligence" could incorporate a physical neural network. One of the primary advantages of a physical neural network is that such a device and applications thereof can be very inexpensive to manufacture, even with present technology. The lithographic techniques required for fabricating the electrodes and channels therebetween has already been perfected and implemented in industry.

Most problems in which a neural network solution is implemented are complex adaptive problems, which change in time. An example is weather prediction. The usefulness of a physical neural network is that it could handle the enormous network needed for such computations and adapt itself in real-time. An example wherein a physical neural network (i.e., Knowm) can be particularly useful is the Personal Digital Assistant (PDA). PDA's are well known in the art. A physical neural network applied to a PDA device can be advantageous because the physical neural network can ideally function with a large network that could constantly adapt itself to the individual user without devouring too much computational time from the PDA. A physical neural network could also be implemented in many industrial applications, such as developing a real-time systems control to the manufacture of various components. This systems control can be adaptable and totally tailored to the particular application, as necessarily it must.

It will be appreciated that variations of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also that various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

What is claimed is:

1. A variable resistor apparatus, comprising:

a plurality of nanoparticles disposed between two terminals, wherein said plurality of nanoparticles acts as an electrical resistance; and

an electric field applied perpendicular to said plurality of nanoparticles across said two terminals resulting in an alignment of said nanoparticles over time and a decrease in said electrical resistance thereby providing a variable resistor apparatus.

2. The apparatus of claim **1** wherein said nanoparticles among said plurality of nanoparticles comprise nanoconductors.

3. The apparatus of claim **2** wherein said nanoconductors comprise nanotubes.

4. The apparatus of claim **2** wherein said nanoconductors comprise nanowires.

5. The apparatus of claim **1** wherein said plurality of nanoparticles are disposed in a solution within a connection gap formed between said two terminals.

6. The apparatus of claim **5** wherein said solution comprises a solvent.

7. The apparatus of claim **5** wherein said solution comprises a suspension of said nanoparticles forming a mixture.

8. The apparatus of claim **5** wherein said solution comprises a liquid.

9. The apparatus of claim **5** wherein said solution comprises a gel.

10. The apparatus of claim **5** wherein said solution comprises a gas.

11. The apparatus of claim **5** wherein said solution comprises a dielectric.

12. A variable resistor apparatus, comprising:

a plurality of nanoparticles disposed in a solution within a connection gap formed between two terminals, wherein said plurality of nanoparticles acts as an electrical resistance; and

an electric field applied to said plurality of nanoparticles across said two terminals, said electric field perpendicular to said plurality of nanoparticles, resulting in an alignment of said nanoparticles over time and a decrease in said electrical resistance thereby providing a variable resistor apparatus.

13. The apparatus of claim **12** wherein said solution comprises a solvent.

14. The apparatus of claim **12** wherein said solution comprises a suspension of said nanoparticles forming a mixture.

15. The apparatus of claim **12** wherein said solution comprises at least one of the following: a liquid, a gel, a gas or a dielectric.

16. A variable resistor apparatus, comprising:

a plurality of nanoconductors disposed between in a solution within a connection gap formed between two terminals, wherein said plurality of nanoconductors acts as an electrical resistance; and

an electric field applied to said plurality of nanoconductors across said two terminals, said electric field perpendicular to said plurality of nanoconductors, resulting in an alignment of said nanoconductors over time and a decrease in said electrical resistance thereby providing a variable resistor apparatus.

17. The apparatus of claim 16 wherein said solution comprises a solvent.

18. The apparatus of claim 16 wherein said solution comprises a suspension of said nanoconductors forming a mixture.

19. The apparatus of claim 16 wherein said solution comprises at least one of the following: a liquid, a gel, a gas or a dielectric.

* * * * *

US006889216B2

(12) United States Patent
Nugent

(10) Patent No.: US 6,889,216 B2
(45) Date of Patent: May 3, 2005

(54) **PHYSICAL NEURAL NETWORK DESIGN INCORPORATING NANOTECHNOLOGY**

(75) Inventor: **Alex Nugent**, Santa Fe, NM (US)

(73) Assignee: **Knowm Tech, LLC**, Albuquerque, NM (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 585 days.

(21) Appl. No.: **10/095,273**

(22) Filed: **Mar. 12, 2002**

(65) **Prior Publication Data**

US 2003/0177450 A1 Sep. 18, 2003

(51) **Int. Cl.**^{7} **G06F 15/18**
(52) **U.S. Cl.** **706/15**
(58) **Field of Search** 706/15

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,802,951	A	2/1989	Clark et al.	156/630
4,974,146	A	11/1990	Works et al.	364/200
4,988,891	A	1/1991	Mashiko	307/201
5,315,162	A	5/1994	McHardy et al.	307/201
5,422,983	A	6/1995	Castelaz et al.	395/24
5,475,794	A	12/1995	Mashiko	395/24
5,589,692	A	12/1996	Reed	257/23
5,649,063	A	7/1997	Bose	395/22
5,706,404	A	1/1998	Colak	395/24
5,717,832	A	2/1998	Steimle et al.	395/24
5,783,840	A	7/1998	Randall et al.	257/24
5,812,993	A	9/1998	Ginosar et al.	706/26
5,904,545	A	5/1999	Smith et al.	438/455
5,951,881	A	9/1999	Rogers et al.	216/41
5,978,782	A	11/1999	Neely	706/16
6,026,358	A	2/2000	Tomabechi	704/232
6,128,214	A	10/2000	Kuekes et al.	365/151
6,248,529	B1	6/2001	Connolly	435/6
6,256,767	B1	7/2001	Kuekes et al.	716/9
6,282,530	B1	8/2001	Huang	706/41
6,294,450	B1	9/2001	Chen et al.	438/597
6,314,019	B1	11/2001	Kuekes et al.	365/151
6,330,553	B1	12/2001	Uchikawa et al.	706/2
6,339,227	B1	1/2002	Ellenbogen	257/40
2001/0004471	A1	6/2001	Zhang	427/372.2
2001/0023986	A1	9/2001	Mancevski	257/741
2001/0024633	A1	9/2001	Lee et al.	423/447.3
2001/0044114	A1	11/2001	Connolly	435/6
2002/0001905	A1	1/2002	Choi et al.	438/268

FOREIGN PATENT DOCUMENTS

EP	1 022 764 A1	1/2000
EP	1 046 613 A2	4/2000
EP	1 100 106 A2	5/2001
EP	1 069 206 A3	7/2001
EP	1 115 135 A1	7/2001
EP	1 134 304 A2	9/2001
WO	WO 00/44094	7/2000

OTHER PUBLICATIONS

Schoenbach et al, "Bioelectrics–New Applications for Pulsed Power Technology", IEEE Digest of Technical Papers on Pulsed Power Plasma Science, Jun. 2001.*

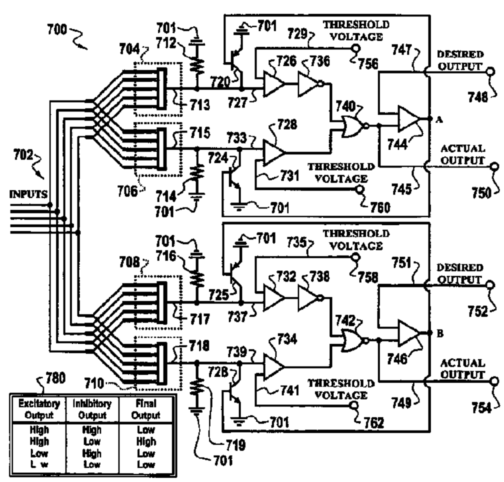
(Continued)

Primary Examiner—George Davis
(74) *Attorney, Agent, or Firm*—Kermit D. Lopez; Luis M. Ortiz; Ortiz & Lopez, PLLC

(57) **ABSTRACT**

A physical neural network based on nanotechnology, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, which are formed from a plurality of interconnected nanoconnections formed from nanoconductors. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes are strengthened or weakened according to an application of an electric field.

52 Claims, 6 Drawing Sheets



OTHER PUBLICATIONS

PCT Notification of Transmittal of the International Search Report or the Declaration, Date of Mailing Jun. 4, 2004.

Peter Weiss, "Circuitry in a Nanowire: Novel Growth Method May Transform Chips," Science News Online, vol. 161, No. 6; Feb. 9, 2002.

Press Release, "Nanowire–based electronics and optics comes one step closer," Eureka Alert, American Chemical Society; Feb. 1, 2002.

Weeks et al., "High–pressure nanolithography using low–energy electrons from a scanning tunneling microscope," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 38–42; Dec. 12, 2001.

CMP Cientifica, "Nanotech: the tiny revolution"; CMP Cientifica, Nov. 2001.

Diehl, et al., "Self–Assembled, Deterministic Carbon Nanotube Wiring Networks," Angew. Chem. Int. Ed. 2002, 41, No. 2; Received Oct. 22, 2001.

G. Pirio, et al., "Fabrication and electrical characteristics of carbon nanotube field emission microcathodes with an integrated gate electrode," Institute of Physics Publishing, Nanotechnology 13 (2002), pp. 1–4, Oct. 2, 2001.

Leslie Smith, "An Introduction to Neural Networks," Center for Cognitive and Computational Neuroscience, Dept. of Computing & Mathematics, Unversity of Stirling, Oct. 25, 1996; http//www.cs.stir.ac.uk/~lss/NNIntro/InvSlides.html.

V. Derycke et al., "Carbon Nanotube Inter– and Intramolecular Logic Gates," American Chemical Society, Nano Letters, XXXX, vol. 0, No. 0, A–D.

Mark K. Anderson, "Mega Steps Toward the Nanochip," Wired News, Apr. 27, 2001.

Collins et al., "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," Science, vol. 292, pp. 706–709, Apr. 27, 2001.

Landman et al., "Metal–Semiconductor Nanocontacts: Silicon Nanowires," Physcial Review Letters, vol. 85, No. 9, Aug. 28, 2000.

John G. Spooner, "Tiny tubes mean big chip advances," Cnet News.com, Tech News First, Apr. 26, 2001.

Jeong–Mi Moon et al., "High–Yield Purification Process of Singewalled Carbon Nanotubes," J. Phys. Chem. B 2001, 105, pp. 5677–5681.

"A New Class of Nanostructure: Semiconducting Nanobelts Offer Potential for Nanosensors and Nanoelectronics," Mar. 12, 2001, http://www.sciencedaily.com/releases/2001/03/010309080953.htm.

Hermanson et al., "Dielectrophoretic Assembly of Electrically Functional Microwires from Nanoparticles Suspensions," Materials Science, vol. 294, No. 5544, Issue of Nov 2, 2001, pp. 1082–1086.

Press Release, "Toshiba Demonstrates Operation of Single–Electron Transistor Circuit at Room Temperature," Toshiba, Jan. 10, 2001.

J. Appenzeller et al., "Optimized contact configuration for the study of transport phenomena in ropes of single–wall carbon nanotubes," Applied Physics Letters, vol. 78, No. 21, pp. 3313–3315, May 21, 2001.

David Rotman, "Molecular Memory, Replacing silicon with organic molecules could mean tiny supercomputers," Technology Review, May 2001, p. 46.

Westervelt et al., "Molecular Electronics,"NSF Functional Nanostructures Grant 9871810, NSF Partnership in Nanotechnology Conference, Jan. 29–30, 2001; http://www.unix-.oit.umass.edu/~nano/NewFiles/FN19_Harvard.pdf.

Niyogi et al., "Chromatographic Purification of Soluble Single–Walled Carbon Nanotubes (s–SWNTs)," J. Am. Chem. Soc 2001, 123, pp. 733–734, Received Jul. 10, 2000.

Duan et al., "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices," Nature, vol. 409, Jan. 4, 2001, pp. 67–69.

Paulson, et al., "Tunable Resistance of a Carbon Nanotube–Graphite Interface," Science, vol. 290, Dec. 1, 2000, pp. 1742–1744.

Wei et al., "Reliability and current carrying capacity of carbon nanotubes," Applied Physics Letters, vol. 79, No. 8, Aug. 20, 2001, pp. 1172–1174.

Collins et al., "Nanotubes for Electronics," Scientific American, Dec. 2000, pp. 62–69.

Avouris et al., "Carbon nanotubes: nanomechanics, manipulation, and electronic devices," Applied Surface Science 141 (1999), pp.201–209.

Smith et al., "Electric–field assisted assembly and aligment of metallic nanowires," Applied Physics Letters, vol. 77, No. 9, Aug. 28, 2000, pp. 1399–1401.

Hone et al., "Electrical and thermal transport properties of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 666–668.

Smith et al., "Structural anisotropy of magnetically aligned single wall carbon nanotube films," Applied Physics Letters, vol. 77, No. 5, Jul. 31, 2000, pp. 663–665.

Andriotis et al., "Various bonding configurations of transition–metal atoms on carbon nanotubes: Their effect on contact resistance," Applied Physics Letters, vol. 76, No. 26, Jun. 26, 2000, pp. 3890–3892.

Chen et al., "Aligning single–wall carbon nanotubes with an alternating–current electric field," Applied Physics Letters, vol. 78, No. 23, Jun. 4, 2001, pp. 3714–3716.

Bezryadin et al., "Self–assembled chains of graphitized carbon nanoparticles," Applied Physics Letters, vol. 74, No. 18, May 3, 1999, pp. 2699–2701.

Bezryadin et al., "Evolution of avalanche conducting states in electrorheological liquids," Physical Review E, vol. 59, No. 6, Jun. 1999, pp. 6896–6901.

Liu et al., "Fullerene Pipes," Science, vol. 280, May 22, 1998, pp. 1253–1255.

Yamamoto et al., "Orientation and purification of carbon nanotubes using ac electrophoresis," J. Phys. D: Appl. Phys 31 (1998) L34–L36.

Bandow et al., "Purification of Single–Wall Carbon Nanotubes by Microfiltration," J. Phys. Chem. B 1997, 101, pp. 8839–8842.

Tohji et al., "Purifying single walled nanotubes," Nature, vol. 383, Oct. 24, 1996, p. 679.

Dejan Rakovic, "Hierarchical Neural Networks and Brainwaves: Towards a Theory of Consciousness," Brain & Consciousness: Proc. ECPD Workshop (ECPD, Belgrade, 1997), pp. 189–204.

Dave Anderson & George McNeill, "Artifical Neural Networks Technology," A DACS (Data & Analysis Center for Software) State–of–the Art Report, Contract No. F30602–89–C–0082, ELIN: A011, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York, Aug. 20, 1992.

Greg Mitchell, "Sub–50 nm Device Fabrication Strategies," Project No. 890–00, Cornell Nanofabrication Facility, Electronics—p. 90–91, National Nanofabrication Users Network.

John–William DeClaris, "An Introduction to Neural Networks," http://www.ee.umd.edu/medlab/neural/nnl.html.
"Neural Networks," StatSoft, Inc., http://www.statsoftinc.com/textbook/stevnet.html.
Stephen Jones, "Neural Networks and the Computation Brain or Maters relating to Artificial Intelligence," The Brain Project, http://www.culture.com.au/brain_proj/neur_net.htm.
David W. Clark, "An Introduction to Neural Networks"; http://members.home.net/neuralnet/introtonn/index.htm.
"A Basic Introduction to Neural Networks"; http://blizzard-.gis.uiuc.edu/htmldocs/Neural/neural.html.
Meyer et al., "Computational neural networks: a general purpose tool for nanotechnology," http://www.foresight.org/Conferences/MNT05/Abstracts/Meyeabst.html.
Saito et al., "A 1M Synapse Self–Learning Digital Neural Network Chip," ISSCC, pp. 6.5–1 to 6.5–10, IEEE 1998.
Espejo, et al., "A 16×16 Cellular Neural Network Chip for Connected Component Detection," Jun. 30 1999; http://www.imse.cnm.csic.es/Chipcat/espeio/chip–2.pdf.
Pati et al., "Neural Networks for Tactile Perception," Systems Research Center and Dept. of Electrical Engineering, University of Maryland and U.S. Naval Research Laboratory. 1987; http://www.isr.umd.edu/TechReports/ISR/1987/TR_87–123/TR_87–123.phtml.
Osamu Fujita, "Statistical estimation of the number of hidden units for feedforward neural networks," Neural Networks 11 (1998), pp. 851–859.

* cited by examiner

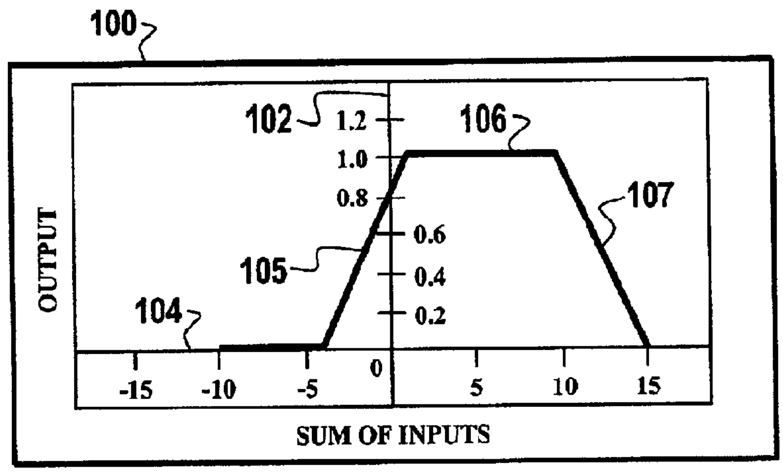


Figure 1

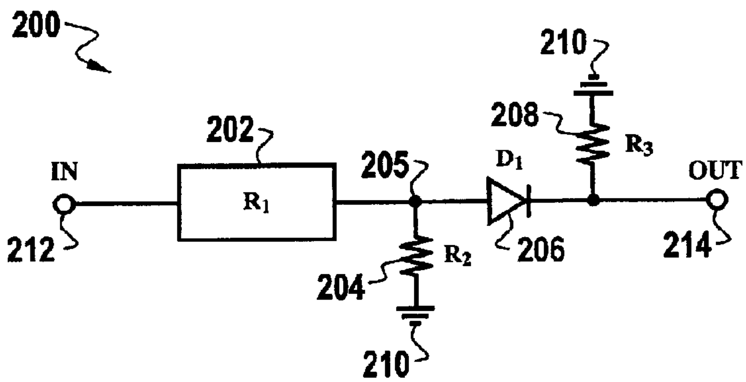


Figure 2

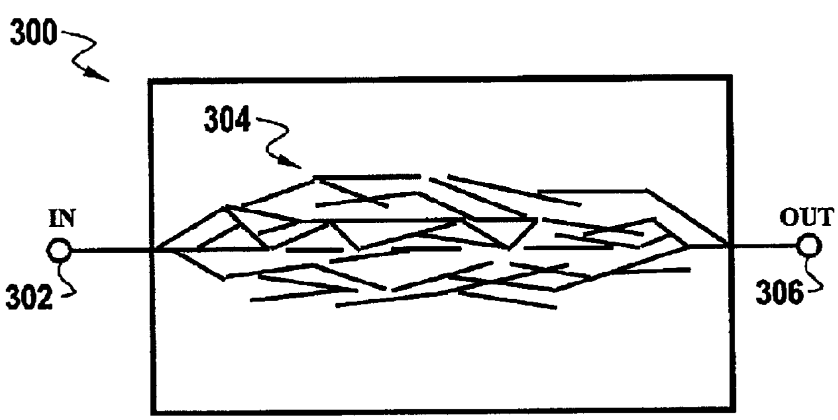


Figure 3

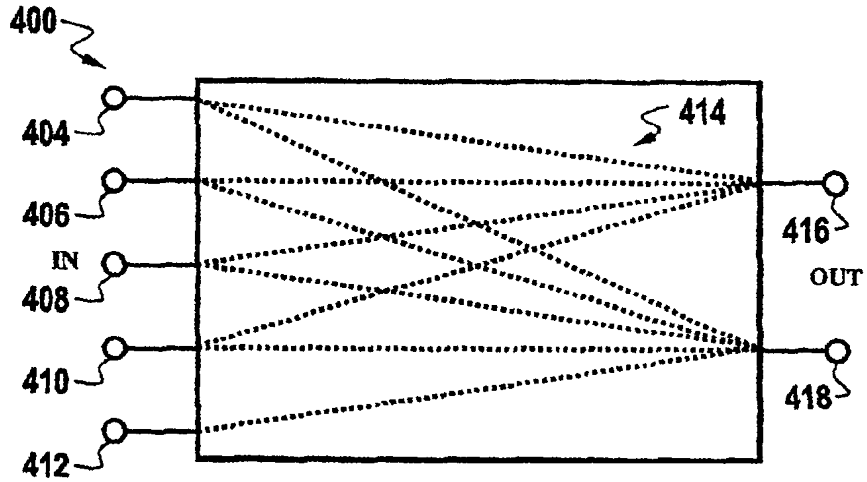


Figure 4

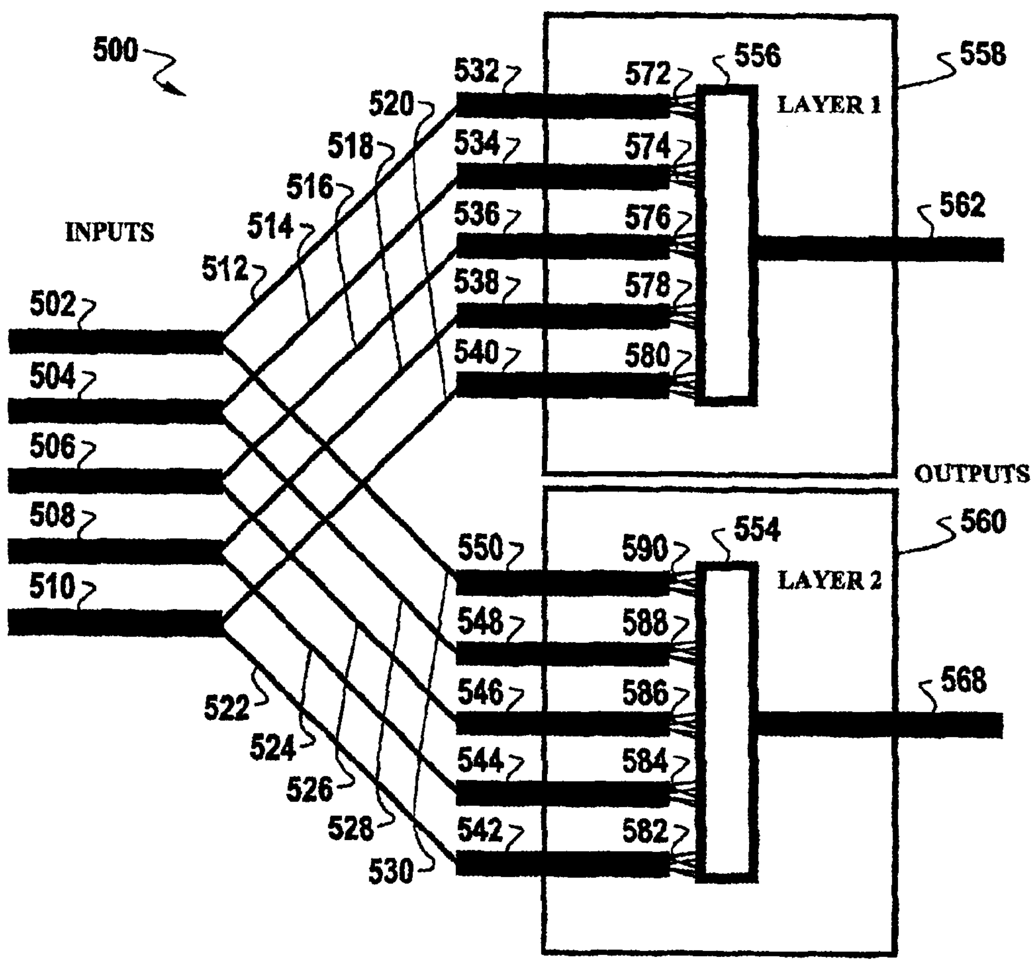


Figure 5

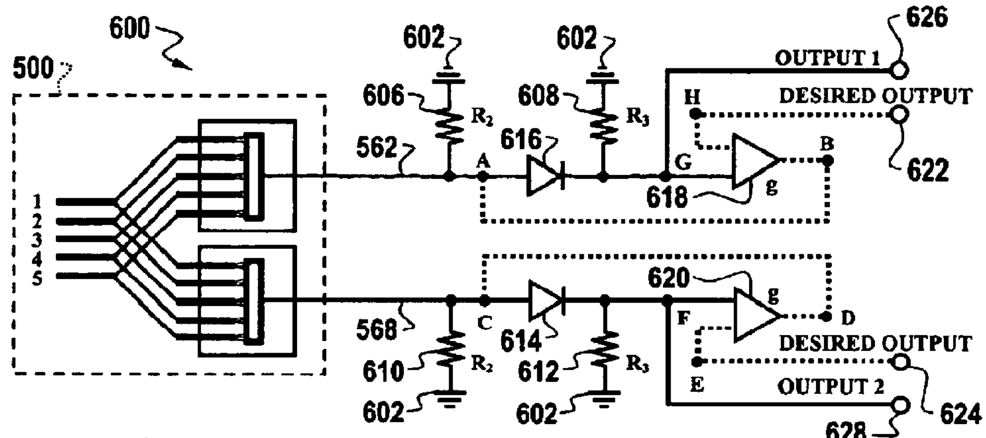
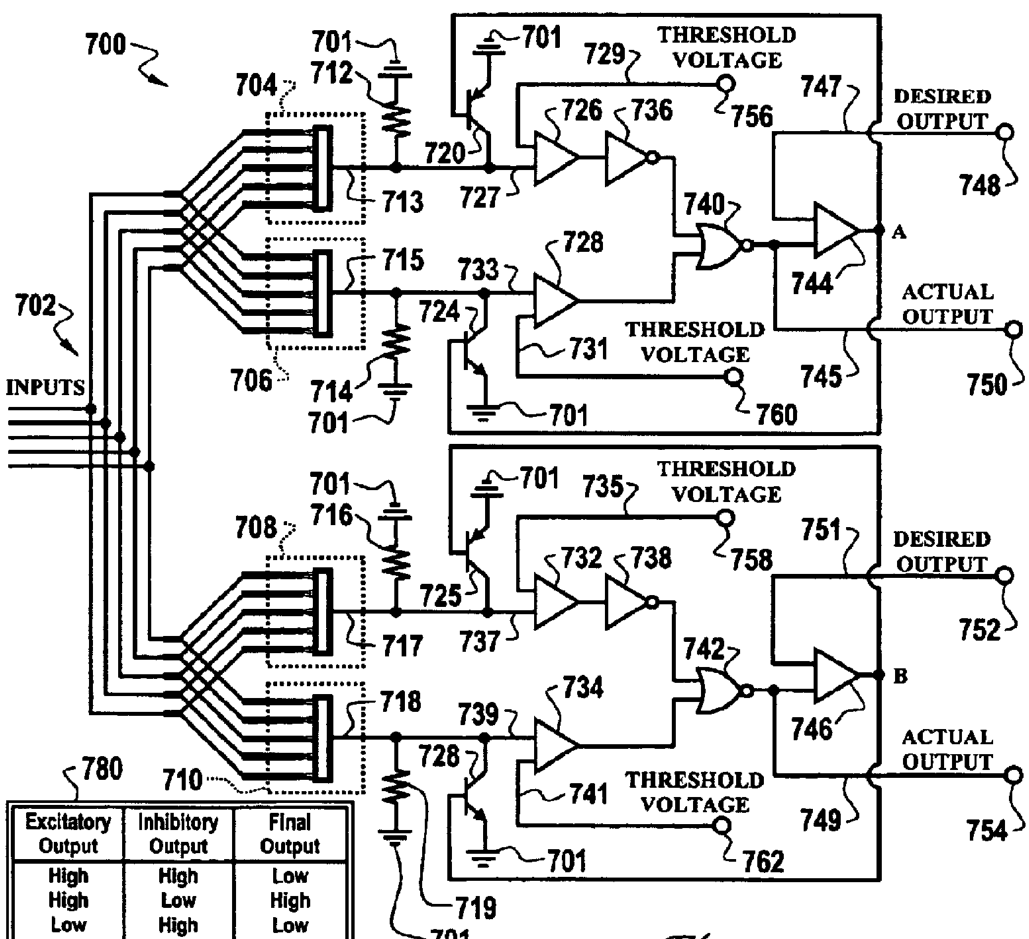


Figure 6



Excitatory Output	Inhibitory Output	Final Output
High	High	Low
High	Low	High
Low	High	Low
L w	Low	Low

Figure 7

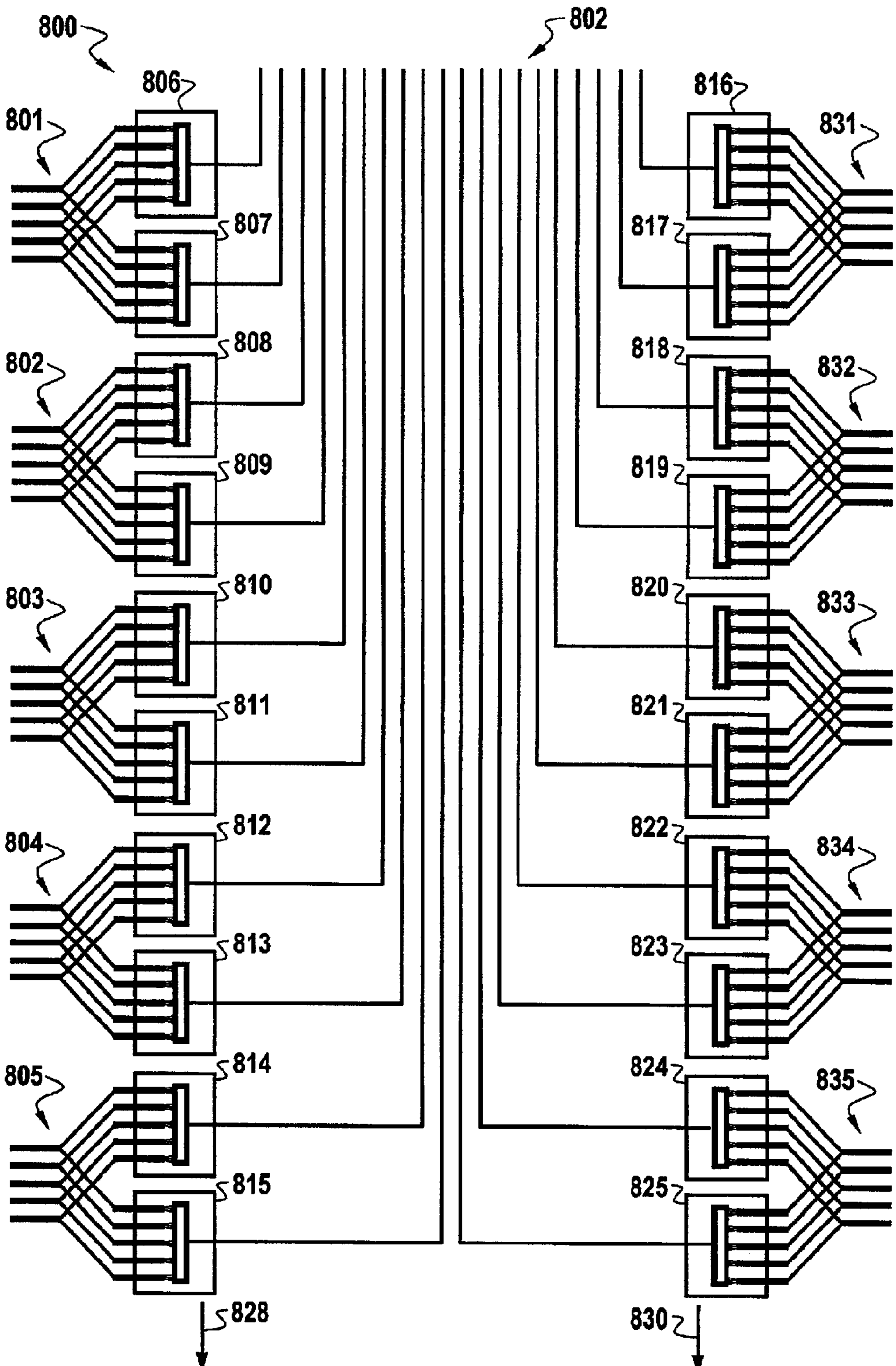

800
802
806
816
801
831
807
817
808
818
802
832
809
819
810
820
803
833
811
821
812
822
804
834
813
823
814
824
805
835
815
825
828
830

Figure 8

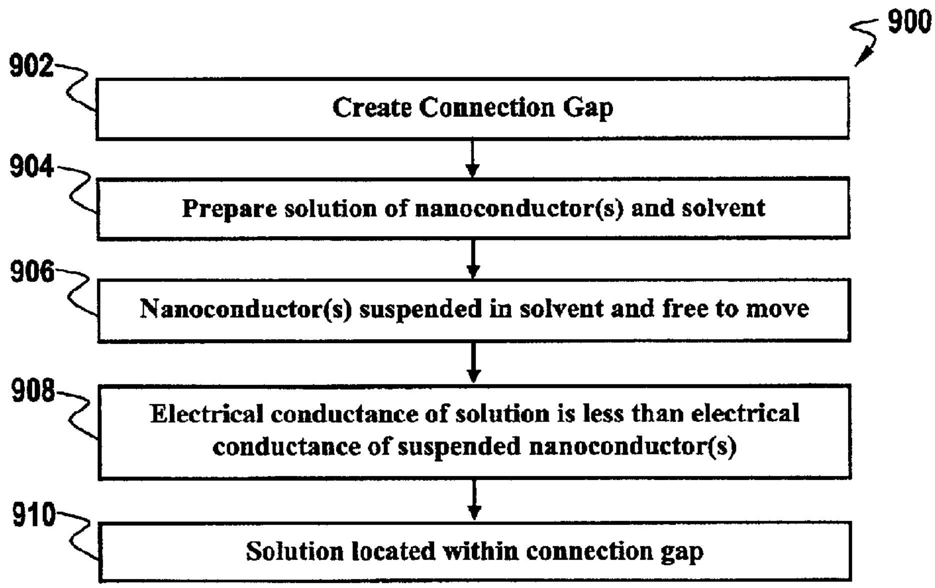


Figure 9

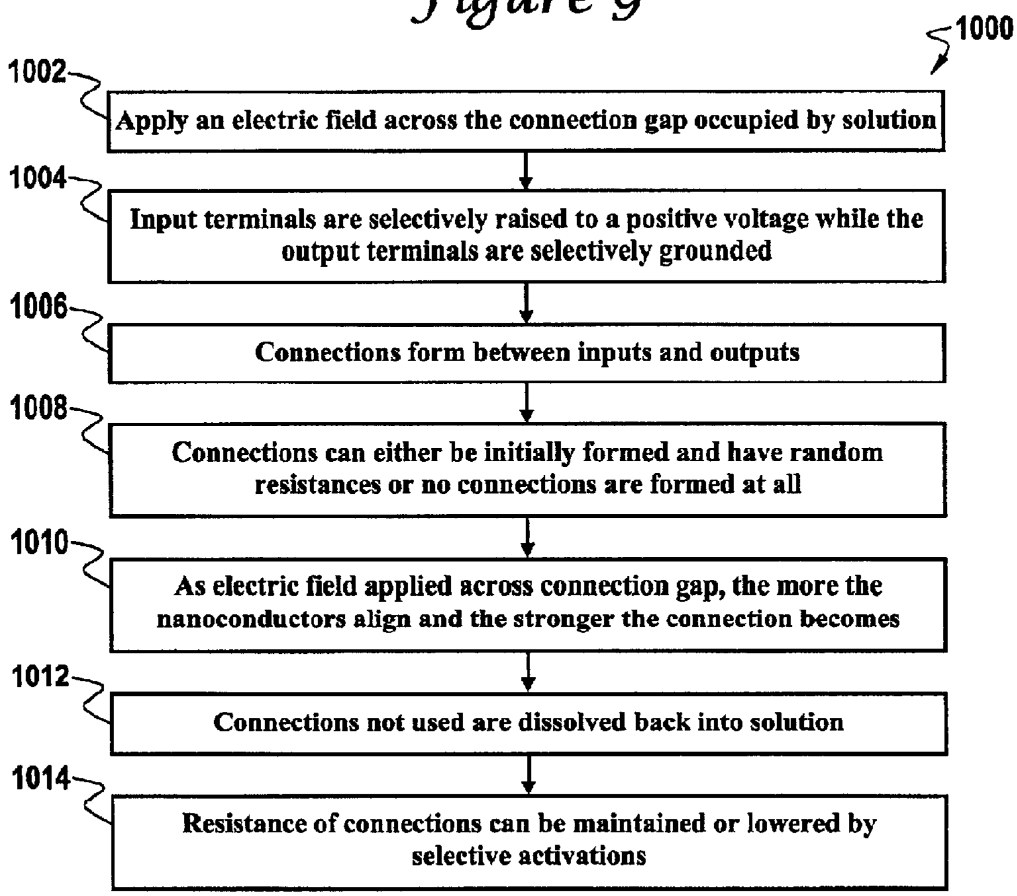


Figure 10

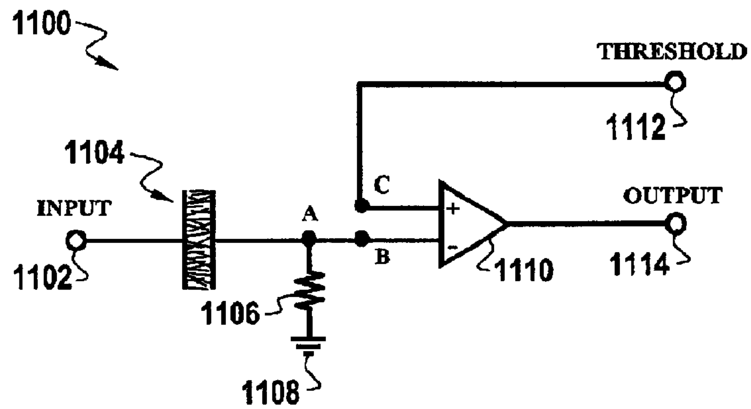


Figure 11

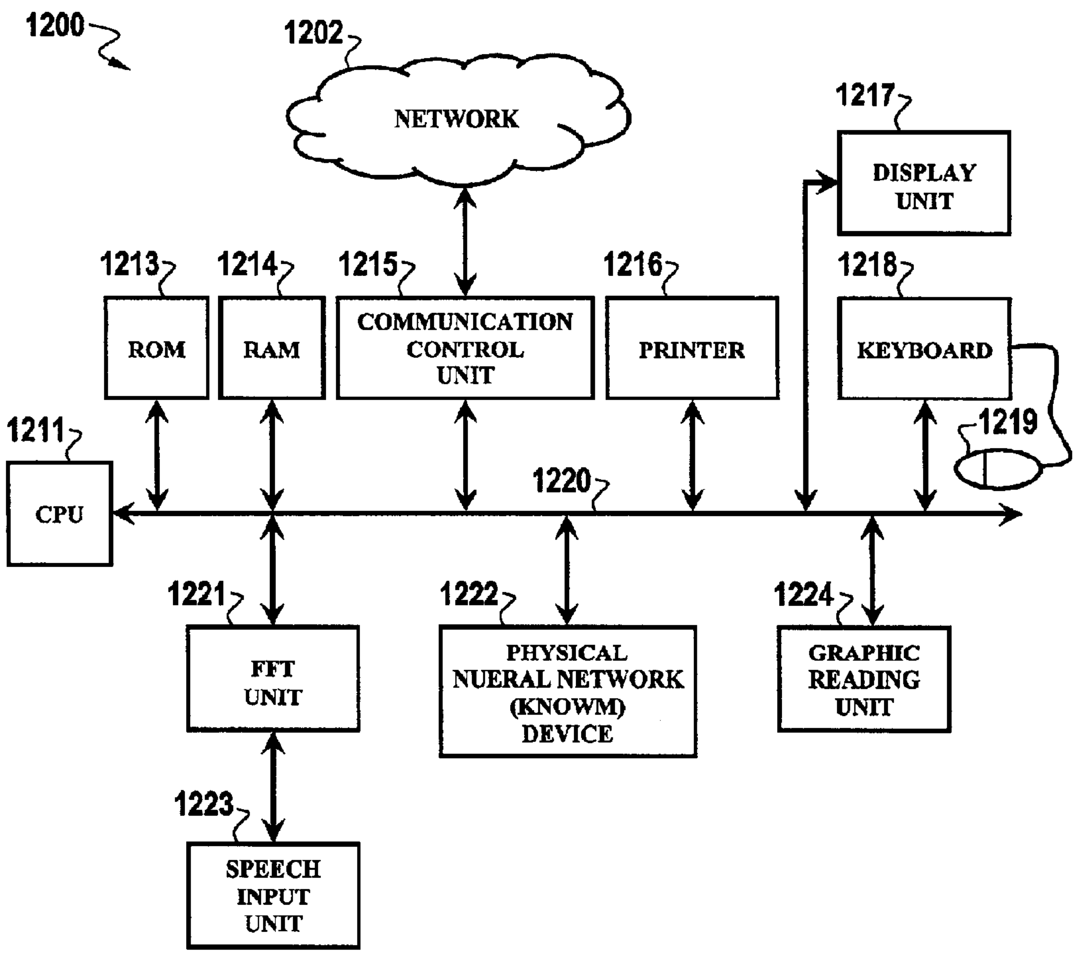


Figure 12

PHYSICAL NEURAL NETWORK DESIGN INCORPORATING NANOTECHNOLOGY

TECHNICAL FIELD

The present invention generally relates to nanotechnology. The present invention also relates to neural networks and neural computing systems and methods thereof. The present invention also relates to physical neural networks, which may be constructed based on nanotechnology. The present invention also related to VLSI (Very Large Scale Integrated) analog neural network chips. The present invention also relates to nanoconductors, such as nanotubes and nanowires.

BACKGROUND OF THE INVENTION

Neural networks are computational systems that permit computers to essentially function in a manner analogous to that of the human brain. Neural networks do not utilize the traditional digital model of manipulating 0's and 1's. Instead, neural networks create connections between processing elements, which are equivalent to neurons of a human brain. Neural networks are thus based on various electronic circuits that are modeled on human nerve cells (i.e., neurons). Generally, a neural network is an information-processing network, which is inspired by the manner in which a human brain performs a particular task or function of interest. Computational or artificial neural networks are thus inspired by biological neural systems. The elementary building block of biological neural systems is of course the neuron, the modifiable connections between the neurons, and the topology of the network.

Biologically inspired artificial neural networks have opened up new possibilities to apply computation to areas that were previously thought to be the exclusive domain of human intelligence. Neural networks learn and remember in ways that resemble human processes. Areas that show the greatest promise for neural networks, such as pattern classification tasks such as speech and image recognition, are areas where conventional computers and data-processing systems have had the greatest difficulty.

In general, artificial neural networks are systems composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural nets. The computational elements, or nodes, are connected via variable weights that are typically adapted during use to improve performance. Thus, in solving a problem, neural net models can explore many competing hypothesis simultaneously using massively parallel nets composed of many computational elements connected by links with variable weights. In contrast, with conventional von Neumann computers, an algorithm must first be developed manually, and a program of instructions written and executed sequentially. In some applications, this has proved extremely difficult. This makes conventional computers unsuitable for many real-time problems. A description and examples of artificial neural networks are disclosed in the publication entitled "Artificial Neural Networks Technology," by Dave Anderson and George McNeill, Aug. 10, 1992, a DACS (Data & Analysis Center for Software) State-of-the-Art Report under Contract Number F30602-89-C-0082, Rome Laboratory RL/C3C, Griffiss Air Force Base, New York.

In a neural network, "neuron-like" nodes can output a signal based on the sum of their inputs, the output being the result of an activation function. In a neural network, there

exists a plurality of connections, which are electrically coupled among a plurality of neurons. The connections serve as communication bridges among of a plurality of neurons coupled thereto. A network of such neuron-like nodes has the ability to process information in a variety of useful ways. By adjusting the connection values between neurons in a network, one can match certain inputs with desired outputs.

One does not program a neural network. Instead, one "teaches" a neural network by examples. Of course, there are many variations. For instance, some networks do not require examples and extract information directly from the input data. The two variations are thus called supervised and unsupervised learning. Neural networks are currently used in applications such as noise filtering, face and voice recognition and pattern recognition. Neural networks can thus be utilized as an advanced mathematical technique for processing information.

Neural networks that have been developed to date are largely software-based. A true neural network (e.g., the human brain) is massively parallel (and therefore very fast computationally) and very adaptable. For example, half of a human brain can suffer a lesion early in its development and not seriously affect its performance. Software simulations are slow because during the learning phase a standard computer must serially calculate connection strengths. When the networks get larger (and therefore more powerful and useful), the computational time becomes enormous. For example, networks with 10,000 connections can easily overwhelm a computer. In comparison, the human brain has about 100 billion neurons, each of which is connected to about 5,000 other neurons. On the other hand, if a network is trained to perform a specific task, perhaps taking many days or months to train, the final useful result can be etched onto a piece of silicon and also mass-produced.

A number of software simulations of neural networks have been developed. Because software simulations are performed on conventional sequential computers, however, they do not take advantage of the inherent parallelism of neural network architectures. Consequently, they are relatively slow. One frequently used measurement of the speed of a neural network processor is the number of interconnections it can perform per second. For example, the fastest software simulations available can perform up to about 18 million interconnects per second. Such speeds, however, currently require expensive super computers to achieve. Even so, 18 million interconnects per second is still too slow to perform many classes of pattern classification tasks in real time. These include radar target classifications, sonar target classification, automatic speaker identification, automatic speech recognition and electro-cardiogram analysis, etc.

The implementation of neural network systems has lagged somewhat behind their theoretical potential due to the difficulties in building neural network hardware. This is primarily because of the large numbers of neurons and weighted connections required. The emulation of even of the simplest biological nervous systems would require neurons and connections numbering in the millions. Due to the difficulties in building such highly interconnected processors, the currently available neural network hardware systems have not approached this level of complexity. Another disadvantage of hardware systems is that they typically are often custom designed and built to implement one particular neural network architecture and are not easily, if at all, reconfigurable to implement different architectures. A true physical neural network chip, for example, has not yet been designed and successfully implemented.

The problem with pure hardware implementation of a neural network with technology as it exists today, is the

inability to physically form a great number of connections and neurons. On-chip learning can exist, but the size of the network would be limited by digital processing methods and associated electronic circuitry. One of the difficulties in creating true physical neural networks lies in the highly complex manner in which a physical neural network must be designed and built. The present inventor believes that solutions to creating a true physical and artificial neural network lies in the use of nanotechnology and the implementation of analog variable connections. The term "Nanotechnology" generally refers to nanometer-scale manufacturing processes, materials and devices, as associated with, for example, nanometer-scale lithography and nanometer-scale information storage. Nanometer-scale components find utility in a wide variety of fields, particularly in the fabrication of microelectrical and microelectromechanical systems (commonly referred to as "MEMS"). Microelectrical nano-sized components include transistors, resistors, capacitors and other nano-integrated circuit components. MEMS devices include, for example, micro-sensors, micro-actuators, micro-instruments, micro-optics, and the like.

In general, nanotechnology presents a solution to the problems faced in the rapid pace of computer chip design in recent years. According to Moore's law, the number of switches that can be produced on a computer chip has doubled every 18 months. Chips now can hold millions of transistors. However, it is becoming increasingly difficult to increase the number of elements on a chip using present technologies. At the present rate, in the next few years the theoretical limit of silicon based chips will be reached. Because the number of elements, which can be manufactured on a chip, determines the data storage and processing capabilities of microchips, new technologies are required which will allow for the development of higher performance chips.

Present chip technology is also limiting when wires need to be crossed on a chip. For the most part, the design of a computer chip is limited to two dimensions. Each time a circuit must cross another circuit, another layer must be added to the chip. This increases the cost and decreases the speed of the resulting chip. A number of alternatives to standard silicon based complementary metal oxide semiconductor ("CMOS") devices have been proposed. The common goal is to produce logic devices on a nanometer scale. Such dimensions are more commonly associated with molecules than integrated circuits.

Integrated circuits and electrical components thereof, which can be produced at a molecular and nanometer scale, include devices such as carbon nanotubes and nanowires, which essentially are nanoscale conductors ("nanoconductors"). Nanocondcutors are tiny conductive tubes (i.e., hollow) or wires (i.e., solid) with a very small size scale (e.g., 1.0–100 nanometers in diameter and hundreds of microns in length). Their structure and fabrication have been widely reported and are well known in the art. Carbon nanotubes, for example, exhibit a unique atomic arrangement, and possess useful physical properties such as one-dimensional electrical behavior, quantum conductance, and ballistic electron transport.

Carbon nanotubes are among the smallest dimensioned nanotube materials with a generally high aspect ratio and small diameter. High-quality single-walled carbon nanotubes can be grown as randomly oriented, needle-like or spaghetti-like tangled tubules. They can be grown by a number of fabrication methods, including chemical vapor deposition (CVD), laser ablation or electric arc growth. Carbon nanotubes can be grown on a substrate by catalytic decomposition of hydrocarbon containing precursors such as ethylene, methane, or benzene. Nucleation layers, such as thin coatings of Ni, Co, or Fe are often intentionally added onto the substrate surface in order to nucleate a multiplicity of isolated nanotubes. Carbon nanotubes can also be nucleated and grown on a substrate without a metal nucleating layer by using a precursor including one or more of these metal atoms. Semiconductor nanowires can be grown on substrates by similar processes.

Attempts have been made to construct electronic devices utilizing nano-sized electrical devices and components. For example, a molecular wire crossbar memory is disclosed in U.S. Pat. No. 6,128,214 entitled "Molecular Wire Crossbar Memory" dated Oct. 3, 2000 to Kuekes et al. Kuekes et al disclose a memory device that is constructed from crossbar arrays of nanowires sandwiching molecules that act as on/off switches. The device is formed from a plurality of nanometer-scale devices, each device comprising a junction formed by a pair of crossed wires where one wire crosses another and at least one connector species connects the pair of crossed wires in the junction. The connector species comprises a bi-stable molecular switch. The junction forms either a resistor or a diode or an asymmetric non-linear resistor. The junction has a state that is capable of being altered by application of a first voltage and sensed by the application of a second, non-destructive voltage. A series of related patents attempts to cover everything from molecular logic to how to chemically assemble these devices.

Such a molecular crossbar device has two general applications. The notion of transistors built from nanotubes and relying on nanotube properties is being pursued. Second, two wires can be selectively brought to a certain voltage and the resulting electrostatic force attracts them. When they touch, the Van der Walls force keeps them in contact with each other and a "bit" is stored. The connections in this apparatus can therefore be utilized for a standard (i.e., binary and serial) computer. The inventors of such a device thus desire to coax a nanoconductor into a binary storage media or a transistor. As it turns out, such a device is easier to utilize as a storage device.

The molecular wire crossbar memory device disclosed in Kuekes et al and related patents thereof simply comprise a digital storage medium that functions at a nano-sized level. Such a device, however, is not well-suited for non-linear and analog functions. Neural networks are non-linear in nature and naturally analog. A neural network is a very non-linear system, in that small changes to its input can create large changes in its output. To date, nanotechnology has not been applied to the creation of truly physical neural networks.

Based on the foregoing, the present inventor believes that a physical neural network which incorporates nanotechnology is a solution to the problems encountered by prior art neural network solutions. In particular, the present inventor believes that a true physical neural network can be designed and constructed without relying on computer simulations for training, or relying on standard digital (binary) memory to store connections strengths.

BRIEF SUMMARY OF THE INVENTION

The following summary of the invention is provided to facilitate an understanding of some of the innovative features unique to the present invention, and is not intended to be a full description. A full appreciation of the various aspects of the invention can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

It is, therefore, one aspect of the present invention to provide a physical neural network.

It is therefore another aspect of the present to provide a physical neural network, which can be formed and implemented utilizing nanotechnology.

It is still another aspect of the present invention to provide a physical neural network, which can be formed from a plurality of interconnected nanoconnections or nanoconnectors.

It is a further aspect of the present invention to provide neuron like nodes, which can be formed and implemented utilizing nanotechnology;

It is also an aspect of the present invention to provide a physical neural network that can be formed from one or more neuron-like nodes.

It is yet a further aspect of the present invention to provide a physical neural network, which can be formed from a plurality of nanoconductors, such as, for example, nanowires and/or nanotubes.

It is still an additional aspect of the present invention to provide a physical neural network, which can be implemented physically in the form of a chip structure.

The above and other aspects can be achieved as is now described. A physical neural network based on nanotechnology is disclosed herein, including methods thereof. Such a physical neural network generally includes one or more neuron-like nodes, connected to a plurality of interconnected nanoconnections. Each neuron-like node sums one or more input signals and generates one or more output signals based on a threshold associated with the input signal. The physical neural network also includes a connection network formed from the interconnected nanoconnections, such that the interconnected nanoconnections used thereof by one or more of the neuron-like nodes can be strengthened or weakened according to an application of an electric field. Alignment has also been observed with a magnetic field, but electric fields are generally more practical. Note that the connection network is associated with one or more of the neuron-like nodes.

The output signal is generally based on a threshold below which the output signal is not generated and above which the output signal is generated. The transition from zero output to high output need not necessarily be abrupt or non linear. The connection network comprises a number of layers of nanoconnections, wherein the number of layers is equal to a number of desired outputs from the connection network. The nanoconnections are formed without influence from disturbances resulting from other nanoconnections thereof. Such nanoconnections may be formed from an electrically conducting material. The electrically conducting material is chosen such that a dipole is induced in the electrically conducting material in the presence of an electric field. Such a nanoconnection may comprise a nanoconductor.

The connection network itself may comprise a connection network structure having a connection gap formed therein, and a solution located within the connection gap, such that the solution comprises a solvent or suspension and one or more nanoconductors. Preferably, a plurality of nanoconductors is present in the solution (i.e., mixture). Note that such a solution may comprise a liquid and/or gas. An electric field can then be applied across the connection gap to permit the alignment of one or more of the nanoconductors within the connection gap. The nanoconductors can be suspended in the solvent, or can lie at the bottom of the connection gap on the surface of the chip. Studies have shown that nanotubes can align both in the suspension and/or on the surface of the gap. The electrical conductance of the mixture is less than the electrical conductance of the nanoconductors within the solution.

The nanoconductors within the connection gap thus experience an increased alignment in accordance with an increase in the electric field applied across the connection gap. Thus, nanoconnections of the neuron-like node that are utilized most frequently by the neuron-like node become stronger with each use thereof. The nanoconnections that are utilized least frequently become increasingly weak and eventually dissolve back into the solution. The nanoconnections may or may not comprise a resistance, which can be raised or lowered by a selective activation of a nanoconnection. They can be configured as nanoconductors such as, for example, a nanotube or nanowire. An example of a nanotube, which may be implemented in accordance with the invention described herein, is a carbon nanotube. Additionally, such nanoconnections may be configured as a negative connection associated with the neuron-like node.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying figures, in which like reference numerals refer to identical or functionally-similar elements throughout the separate views and which are incorporated in and form part of the specification, further illustrate the present invention and, together with the detailed description of the invention, serve to explain the principles of the present invention.

FIG. **1** illustrates a graph illustrating a typical activation function that can be implemented in accordance with the physical neural network of the present invention;

FIG. **2** depicts a schematic diagram illustrating a diode configuration as a neuron, in accordance with a preferred embodiment of the present invention;

FIG. **3** illustrates a block diagram illustrating a network of nanoconnections formed between two electrodes, in accordance with a preferred embodiment of the present invention;

FIG. **4** depicts a block diagram illustrating a plurality of connections between inputs and outputs of a physical neural network, in accordance with a preferred embodiment of the present invention;

FIG. **5** illustrates a schematic diagram of a physical neural network that can be created without disturbances, in accordance with a preferred embodiment of the present invention;

FIG. **6** depicts a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment of the present invention;

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network that can be implemented in accordance with an alternative embodiment of the present invention;

FIG. **8** depicts a schematic diagram of a chip layout for a connection network that may be implemented in accordance with an alternative embodiment of the present invention;

FIG. **9** illustrates a flow chart of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment of the present invention;

FIG. **10** depicts a flow chart of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment of the present invention;

FIG. **11** illustrates a schematic diagram of a circuit illustrating temporal summation within a neuron, in accordance with a preferred embodiment of the present invention; and

FIG. **12** depicts a block diagram illustrating a pattern recognition system, which may be implemented with a

physical neural network device, in accordance with an alternative embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate an embodiment of the present invention and are not intended to limit the scope of the invention.

The physical neural network described and disclosed herein is different from prior art forms of neural networks in that the disclosed physical neural network does not require a computer simulation for training, nor is its architecture based on any current neural network hardware device. The design of the physical neural network of the present invention is actually quite "organic". The physical neural network described herein is generally fast and adaptable, no matter how large such a physical neural network becomes. The physical neural network described herein can be referred to generically as a Knowm. The terms "physical neural network" and "Knowm" can be utilized interchangeably to refer to the same device, network, or structure.

Network orders of magnitude larger than current VSLI neural networks can be built and trained with a standard computer. One consideration for a Knowm is that it must be large enough for its inherent parallelism to shine through. Because the connection strengths of such a physical neural network are dependant on the physical movement of nanoconnections thereof, the rate at which a small network can learn is generally very small and a comparable network simulation on a standard computer can be very fast. On the other hand, as the size of the network increases, the time to train the device does not change. Thus, even if the network takes a full second to change a connection value a small amount, if it does the same to a billion connections simultaneously, then its parallel nature begins to express itself.

A physical neural network (i.e., a Knowm) must have two components to function properly. First, the physical neural network must have one or more neuron-like nodes that sum a signal and output a signal based on the amount of input signal received. Such a neuron-like node is generally non-linear in its output. In other words, there should be a certain threshold for input signals, below which nothing is output and above which a constant or nearly constant output is generated or allowed to pass. This is a very basic requirement of standard software-based neural networks, and can be accomplished by an activation function. The second requirement of a physical neural network is the inclusion of a connection network composed of a plurality of interconnected connections (i.e., nanoconnections). Such a connection network is described in greater detail herein.

FIG. 1 illustrates a graph **100** illustrating a typical activation function that can be implemented in accordance with the physical neural network of the present invention. Note that the activation function need not be non-linear, although non-linearity is generally desired for learning complicated input-output relationships. The activation function depicted in FIG. 1 comprises a linear function, and is shown as such for general edification and illustrative purposes only. As explained previously, an activation function may also be non-linear.

As illustrated in FIG. 1, graph **100** includes a horizontal axis **104** representing a sum of inputs, and a vertical axis **102** representing output values. A graphical line **106** indicates threshold values along a range of inputs from approximately −10 to +10 and a range of output values from approximately 0 to 1. As more neural networks (i.e., active inputs) are established, the overall output as indicated at line **105** climbs until the saturation level indicated by line **106** is attained. If a connection is not utilized, then the level of output (i.e., connection strength) begins to fade until it is revived. This phenomenon is analogous to short term memory loss of a human brain. Note that graph **100** is presented for generally illustrative and edification purposes only and is not considered a limiting feature of the present invention.

In a Knowm, the neuron-like node can be configured as a standard diode-based circuit, the diode being the most basic semiconductor electrical component, and the signal it sums may be a voltage. An example of such an arrangement of circuitry is illustrated in FIG. **2**, which generally depicts a schematic diagram illustrating a diode-based configuration as a neuron **200**, in accordance with a preferred embodiment of the present invention. Those skilled in the art can appreciate that the use of such a diode-based configuration is not considered a limiting feature of the present invention, but merely represents one potential arrangement in which the present invention may be implemented.

Although a diode may not necessarily be utilized, its current versus voltage characteristics are non-linear when used with associated resistors and similar to the relationship depicted in FIG. **1**. The use of a diode as a neuron is thus not a limiting feature of the present invention, but is only referenced herein with respect to a preferred embodiment. The use of a diode and associated resistors with respect to a preferred embodiment simply represents one potential "neuron" implementation. Such a configuration can be said to comprise an artificial neuron. It is anticipated that other devices and components may be utilized instead of a diode to construct a physical neural network and a neuron-like node (i.e., artificial neuron), as indicated here.

Thus, neuron **200** comprises a neuron-like node that may include a diode **206**, which is labeled D_1, and a resistor **204**, which is labeled R_2. Resistor **204** is connected to a ground **210** and an input **205** of diode **206**. Additionally, a resistor **202**, which is represented as a block and labeled R_1 can be connected to input **205** of diode **206**. Block **202** includes an input **212**, which comprises an input to neuron **200**. A resistor **208**, which is labeled R_3, is also connected to an output **214** of diode **206**. Additionally, resistor **208** is coupled to ground **210**. Diode **206** in a physical neural network is analogous to a neuron of a human brain, while an associated connection formed thereof, as explained in greater detail herein, is analogous to a synapse of a human brain.

As depicted in FIG. **2**, the output **214** is determined by the connection strength of R_1 (i.e., resistor **202**). If the strength of R_1's connection increases (i.e., the resistance decreases), then the output voltage at output **214** also increases. Because diode **206** conducts essentially no current until its threshold voltage (e.g., approximately 0.6V for silicon) is attained, the output voltage will remain at zero until R_1 conducts enough current to raise the pre-diode voltage to approximately 0.6V. After 0.6V has been achieved, the output voltage at output **214** will increase linearly. Simply adding extra diodes in series or utilizing different diode types may increase the threshold voltage.

An amplifier may also be added to the output **214** of diode **206** so that the output voltage immediately saturates at the diode threshold voltage, thus resembling a step function, until a threshold value and a constant value above the threshold is attained. R_3 (i.e., resistor **208**) functions gener-

ally as a bias for diode **206** (i.e., D_1) and should generally be about 10 times larger than resistor **204** (i.e., R_2). In the circuit configuration illustrated in FIG. **2**, R_1 can actually be configured as a network of connections composed of many inter-connected conducting nanowires (i.e., see FIG. **3**). As explained previously, such connections are analogous to the synapses of a human brain.

FIG. **3** illustrates a block diagram illustrating a network **300** of nanoconnections **304** formed between two electrodes, in accordance with a preferred embodiment of the present invention. Nanoconnections **304** (e.g., nanoconductors) depicted in FIG. **3** are generally located between input **302** and output **306**. The network of nanoconnections depicted in FIG. **3** can be implemented as a network of nanoconductors. Examples of nanoconductors include devices such as, for example, nanowires, nanotubes, and nanoparticles. Nanoconnections **304**, which are analogous to the synapses of a human brain, should be composed of electrical conducting material (i.e., nanoconductors). It should be appreciated by those skilled in the art that such nanoconductors can be provided in a variety of shapes and sizes without departing from the teachings herein.

For example, carbon particles (e.g., granules or bearings) may be used for developing nanoconnections. The nanoconductors utilized to form a connection network may be formed as a plurality of nanoparticles. For example, each nanoconnection within a connection network may be formed from a chain of carbon nanoparticles. In "Self-assembled chains of graphitized carbon nanoparticles" by Bezryadin et al., Applied Physics Letters, Vol. 74, No. 18, pp. 2699–2701, May 3, 1999, for example, a technique is reported, which permits the self-assembly of conducting nanoparticles into long continuous chains. Bezryadin et al. suggest that new approaches could be developed in order to organize nanoparticles into useful electronic devices. Thus, nanoconductors utilized to form a physical neural network (i.e., Knowm) could be formed from nanoparticles.

It should be appreciated by those skilled in the art that the Bezyadin et al reference does not, of course, comprise limiting features of the present invention, nor does it teach, suggest nor anticipate a physical neural network. Rather, such a reference merely demonstrate recent advances in the carbon nanotechnology arts and how such advances may be adapted for use in association with the Knowm-based system described herein. It can be further appreciated that a connection network as disclosed herein may be composed from a variety of different types of nanoconductors. For example, a connection network may be formed from a plurality of nanoconductors, including nanowires, nanotubes and/or nanoparticles. Note that such nanowires, nanotubes and/or nanoparticles, along with other types of nanoconductors can be formed from materials such as carbon or silicon. For example, carbon nanotubes may comprise a type of nanotube that can be utilized in accordance with the present invention.

As illustrated in FIG. **3**, nanoconnections **304** comprise a plurality of interconnected nanoconnections, which from this point forward, can be referred to generally as a "connection network." An individual nanoconnection may constitute a nanoconductor such as, for example, a nanowire, a nanotube, nanoparticles(s), or any other nanoconducting structures. Nanoconnections **304** may comprise a plurality of interconnected nanotubes and/or a plurality of interconnected nanowires. Similarly, nanoconnections **304** may be formed from a plurality of interconnected nanoparticles. A connection network is thus not one connection between two electrodes, but a plurality of connections between inputs and outputs. Nanotubes, nanowires, nanoparticles and/or other nanoconducting structures may be utilized, of course, to construct nanoconnections **304** between input **302** and input **306**. Although a single input **302** and a single input **306** is depicted in FIG. **3**, it can be appreciated that a plurality of inputs and a plurality of outputs may be implemented in accordance with the present invention, rather than simply a single input **302** or a single output **306**.

FIG. **4** depicts a block diagram illustrating a plurality of nanoconnections **414** between inputs **404, 406, 408, 410, 412** and outputs **416** and **418** of a physical neural network, in accordance with a preferred embodiment of the present invention. Inputs **404, 406, 408, 410,** and **412** can provide input signals to connections **414**. Output signals can then be generated from connections **414** via outputs **416** and **418**. A connection network can therefore be configured from the plurality of connections **414**. Such a connection network is generally associated with one or more neuron-like nodes.

The connection network also comprises a plurality of interconnected nanoconnections, wherein each nanoconnection thereof is strengthened or weakened according to an application of an electric field. A connection network is not possible if built in one layer because the presence of one connection can alter the electric field so that other connections between adjacent electrodes could not be formed. Instead, such a connection network can be built in layers, so that each connection thereof can be formed without being influenced by field disturbances resulting from other connections. This can be seen in FIG. **5**.

FIG. **5** illustrates a schematic diagram of a physical neural network **500** that can be created without disturbances, in accordance with a preferred embodiment of the present invention. Physical neural network **500** is composed of a first layer **558** and a second layer **560**. A plurality of inputs **502, 504, 506, 508,** and **510** are respectively provided to layers **558** and **560** respectively via a plurality of input lines **512, 514, 516, 518,** and **520** and a plurality of input lines **522, 524, 526, 528,** and **530**. Input lines **512, 514, 516, 518,** and **520** are further coupled to input lines **532, 534, 536, 538,** and **540** such that each line **532, 534, 536, 538,** and **540** is respectively coupled to nanoconnections **572, 574, 576, 578,** and **580**. Thus, input line **532** is connected to nanconnections **572**. Input line **534** is connected to nanoconnections **574**, and input line **536** is connected to nanoconnections **576**. Similarly, input line **538** is connected to nanconnections **578**, and input line **540** is connected to nanoconnections **580**.

Nanconnections **572, 574, 576, 578,** and **580** may comprise nanoconductors such as, for example, nanotubes and/or nanowires. Nanoconnections **572, 574, 576, 578,** and **580** thus comprise one or more nanoconductors. Additionally, input lines **522, 524, 526, 528,** and **530** are respectively coupled to a plurality of input lines **542, 544, 546, 548** and **550**, which are in turn each respectively coupled to nanoconnections **582, 584, 586, 588,** and **590**. Thus, for example, input line **542** is connected to nanoconnections **582**, while input line **544** is connected to nanoconnections **584**. Similarly, input line **546** is connected to nanoconnections **586** and input line **548** is connected to nanoconnections **588**. Additionally, input line **550** is connected to nanconnections **590**. Box **556** and **554** generally represent simply the output and are thus illustrated connected to outputs **562** and **568**. In other words, outputs **556** and **554** respectively comprise outputs **562** and **568**. The aforementioned input lines and associated components thereof actually comprise physical electronic components, including conducting input and output lines and physical nanoconnections, such as nanotubes and/or nanowires.

Thus, the number of layers **558** and **560** equals the number of desired outputs **562** and **568** from physical neural network **500**. In the previous two figures, every input was potentially connected to every output, but many other configurations are possible. The connection network can be made of any electrically conducting material, although the physics of it requires that they be very small so that they will align with a practical voltage. Carbon nanotubes or any conductive nanowire can be implemented in accordance with the physical neural network described herein.

Such components can form connections between electrodes by the presence of an electric field. For example, the orientation and purification of carbon nanotubes has been demonstrated using ac electrophoresis in isopropyl alcohol, as indicated in "Orientation and purification of carbon nanotubes using ac electrophoresis" by Yamamoto et al., J. Phys. D: Applied Physics, 31(1998), L34–36. Additionally, an electric-field assisted assembly technique used to position individual nanowires suspended in an electric medium between two electrodes defined lithographically on an SiO2 substrate is indicated in "Electric-field assisted assembly and alignment of metallic nanowires," by Smith et al., Applied Physics Letters, Vol. 77, Num. 9, Aug. 28, 200. It can be appreciated by those skilled in the art that such references are not considered limiting features of the present invention, nor do such references teach, suggest or anticipate a physical neural network as described herein. Such references are discussed herein for general background and illustrative purposes only.

Additionally, it has been reported that it is possible to fabricate deterministic wiring networks from single-walled carbon nanotubes (SWNTs) as indicated in "Self-Assembled, Deterministic Carbon Nanotube Wiring Networks" by Diehi, et al. in Angew. Chem. Int. Ed. 2002, 41. No. 2. In addition, the publication "Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices" by Duan, et al., Nature, Vol. 409, Jan. 4, 2001, reports that an electric-field-directed assembly can be used to create highly integrated device arrays from nanowire building blocks. It should be appreciated by those skilled in the art these references do not comprise limiting features of the present invention, nor do such references teach or anticipate a physical neural network. Rather, such references are incorporated herein by reference to demonstrate recent advances in the carbon nanotechnology arts and how such advances may be adapted for use in association with the physical neural network described herein.

The only general requirements for the conducting material utilized to configure the nanoconductors are that such conducting material must conduct electricity, and a dipole should preferably be induced in the material when in the presence of an electric field. Alternatively, the nanoconductors utilized in association with the physical neural network described herein can be configured to include a permanent dipole that is produced by a chemical means, rather than a dipole that is induced by an electric field.

Therefore, it should be appreciated by those skilled in the art that a connection network could also be comprised of other conductive particles that may be developed or found useful in the nanotechnology arts. For example, carbon particles (or "dust") may also be used as nanoconductors in place of nanowires or nanotubes. Such particles may include bearings or granule-like particles.

A connection network can be constructed as follows: A voltage is applied across a gap that is filled with a mixture of nanowires and a "solvent". This mixture could be made of many things. The only requirements are that the conducting wires must be suspended in the solvent, either dissolved or in some sort of suspension, free to move around; the electrical conductance of the substance must be less than the electrical conductance of the suspended conducting wire; and the viscosity of the substance should not be too much so that the conducting wire cannot move when an electric field is applied.

The goal for such a connection network is to develop a network of connections of just the right values so as to satisfy the particular signal-processing requirement—exactly what a neural network does. Such a connection network can be constructed by applying a voltage across a space occupied by the mixture mentioned. To create the connection network, the input terminals are selectively raised to a positive voltage while the output terminals are selectively grounded. Thus, connections can gradually form between the inputs and outputs. The important requirement that makes the physical neural network of the present invention functional as a neural network is that the longer this electric field is applied across a connection gap, or the greater the frequency or amplitude, the more nanotubes and/or nanowires and/or particles align and the stronger the connection thereof becomes. Thus, the connections that are utilized most frequently by the physical neural network become the strongest.

The connections can either be initially formed and have random resistances or no connections may be formed at all. By initially forming random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up from scratch. Depending on the rate of connection decay, having initial random connections could prove faster, although not necessarily. The connection network can adapt itself to the requirements of a given situation regardless of the initial state of the connections. Either initial condition will work, as connections that are not used will "dissolve" back into solution. The resistance of the connection can be maintained or lowered by selective activations of the connection. In other words, if the connection is not used, it will fade away, analogous to the connections between neurons in a human brain. The temperature of the solution can also be maintained at a particular value so that the rate that connections fade away can be controlled. Additionally an electric field can be applied perpendicular to the connections to weaken them, or even erase them out altogether (i.e., as in clear, zero, or reformatting of a "disk").

The nanoconnections may or may not be arranged in an orderly array pattern. The nanoconnections (e.g., nanotubes, nanowires, etc) of a physical neural network do not have to order themselves into neatly formed arrays. They simply float in the solution, or lie at the bottom of the gap, and more or less line up in the presence an electric field. Precise patterns are thus not necessary. In fact, neat and precise patterns may not be desired. Rather, due to the non-linear nature of neural networks, precise patterns could be a drawback rather than an advantage. In fact, it may be desirable that the connections themselves function as poor conductors, so that variable connections are formed thereof, overcoming simply an "on" and "off" structure, which is commonly associated with binary and serial networks and structures thereof.

FIG. **6** depicts a schematic diagram illustrating an example of a physical neural network **600** that can be implemented in accordance an alternative embodiment of the present invention. Note that in FIGS. **5** and **6**, like parts are indicated by like reference numerals. Thus, physical

neural network **600** can be configured, based on physical neural network **500** illustrated in FIG. **5**. In FIG. **6**, inputs **1**, **2**, **3**, **4**, and **5** are indicated, which are respectively analogous to inputs **502**, **504**, **506**, **508**, and **510** illustrated in FIG. **5**. Outputs **562** and **568** are provided to a plurality of electrical components to create a first output **626** (i.e., Output **1**) and a second output **628** (i.e., Output **2**). Output **562** is tied to a resistor **606**, which is labeled R2 and a diode **616** at node A. Output **568** is tied to a resistor **610**, which is also labeled R2 and a diode **614** at node C. Resistors **606** and **610** are each tied to a ground **602**.

Diode **616** is further coupled to a resistor **608**, which is labeled R3, and first output **626**. Additionally, resistor **608** is coupled to ground **602** and an input to an amplifier **618**. An output from amplifier **618**, as indicated at node B and dashed lines thereof, can be tied back to node A. A desired output **622** from amplifier **618** is coupled to amplifier **618** at node H. Diode **614** is coupled to a resistor **612** at node F. Note that resistor **612** is labeled R3. Node F is in turn coupled to an input of amplifier **620** and to second output **628** (i.e., Output **2**). Diode **614** is also connected to second output **628** and an input to amplifier **620** at second output **628**. Note that second output **628** is connected to the input to amplifier **620** at node F. An output from amplifier **620** is further coupled to node D, which in turn is connected to node C. A desired output **624**, which is indicated by a dashed line in FIG. **6**, is also coupled to an input of amplifier **620** at node E.

In FIG. **6**, the training of physical neural network **600** can be accomplished utilizing, for example, op-amp devices (e.g., amplifiers **618** and **620**). By comparing an output (e.g., first output **626**) of physical neural network **600** with a desired output (e.g., desired output **622**), the amplifier (e.g., amplifier **618**) can provide feedback and selectively strengthen connections thereof. For instance, suppose it is desired to output a voltage of +V at first output **626** (i.e., Output **1**) when inputs **1** and **4** are high. When inputs **1** and **4** are taken high, also assume that first output **626** is zero. Amplifier **618** can then compare the desired output (+V) with the actual output (0) and output −V. In this case, −V is equivalent to ground.

The op-amp outputs and grounds the pre-diode junction (i.e., see node A) and causes a greater electric field across inputs **1** and **4** and the layer **1** output. This increased electric field (larger voltage drop) can cause the nanoconductors in the solution between the electrode junctions to align themselves, aggregate, and form a stronger connection between the **1** and **4** electrodes. Feedback can continue to be applied until output of physical neural network **600** matches the desired output. The same procedure can be applied to every output.

In accordance with the aforementioned example, assume that Output **1** was higher than the desired output (i.e., desired output **622**). If this were the case, the op-amp output can be +V and the connection between inputs **1** and **4** and layer one output can be raised to +V. Columbic repulsions between the nanoconductors can force the connection apart, thereby weakening the connection. The feedback will then continue until the desired output is obtained. This is just one training mechanism. One can see that the training mechanism does not require any computations, because it is a simple feedback mechanism.

Such a training mechanism, however, may be implemented in many different forms. Basically, the connections in a connection network must be able to change in accordance with the feedback provided. In other words, the very general notion of connections being strengthened or connections being weakened in a physical system is the essence of a physical neural network (i.e., Knowm). Thus, it can be appreciated that the training of such a physical neural network may not require a "CPU" to calculate connection values thereof. The Knowm can adapt itself. Complicated neural network solutions could be implemented very rapidly "on the fly", much like a human brain adapts as it performs.

The physical neural network disclosed herein thus has a number of broad applications. The core concept of a Knowm, however, is basic. The very basic idea that the connection values between electrode junctions by nanoconductors can be used in a neural network devise is all that required to develop an enormous number of possible configurations and applications thereof.

Another important feature of a physical neural network is the ability to form negative connections. This is an important feature that makes possible inhibitory effects useful in data processing. The basic idea is that the presence of one input can inhibit the effect of another input. In artificial neural networks as they currently exist, this is accomplished by multiplying the input by a negative connection value. Unfortunately, with a Knowm-based device, the connection may only take on zero or positive values under such a scenario.

In other words, either there can be a connection or no connection. A connection can simulate a negative connection by dedicating a particular connection to be negative, but one connection cannot begin positive and through a learning process change to a negative connection. In general, if it starts positive, it can only go to zero. In essence, it is the idea of possessing a negative connection initially that results in the simulation, because this does not occur in a human brain. Only one type of signal generally travels through axon/dendrites in a human brain. That signal is transferred into the flow of a neurotransmitter whose effect on the postsynaptic neuron can be either excitatory or inhibitory, depending on the neuron.

One method for solving this problem is to utilize two sets of connections for the same output, having one set represent the positive connections and the other set represent the negative connections. The output of these two layers can be compared, and the layer with the greater output will output either a high signal or a low signal, depending on the type of connection set (inhibitory or excitatory). This can be seen in FIG. **7**.

FIG. **7** illustrates a schematic diagram illustrating an example of a physical neural network **700** that can be implemented in accordance with an alternative embodiment of the present invention. Physical neural network **700** thus comprises a plurality of inputs **702** (not necessarily binary) which are respectively fed to layers **704**, **706**, **708**, and **710**. Each layer is analogous to the layers depicted earlier, such as for example layers **558** and **560** of FIG. **5**. An output **713** of layer **704** can be connected to a resistor **712**, a transistor **720** and a first input **727** of amplifier **726**. Transistor **720** is generally coupled between ground **701** and first input **727** of amplifier **726**. Resistor **712** is connected to a ground **701**. Note that ground **701** is analogous to ground **602** illustrated in FIG. **6** and ground **210** depicted in FIG. **2**. A second input **729** of amplifier **726** can be connected to a threshold voltage 756. The output of amplifier **726** can in turn be fed to an inverting amplifier **736**.

The output of inverting amplifier **736** can then be input to a NOR device **740**. Similarly, an output **716** of layer **706** may be connected to resistor **714**, transistor **733** and a first input **733** of an amplifier **728**. A threshold voltage 760 is

connected to a second input **737** of amplifier **728**. Resistor **714** is generally coupled between ground **701** and first input **733** of amplifier **728**. Note that first input **733** of amplifier **728** is also generally connected to an output **715** of layer **706**. The output of amplifier **728** can in turn be provided to NOR device **740**. The output from NOR device **740** is generally connected to a first input **745** of an amplifier **744**. An actual output **750** can be taken from first input **745** to amplifier **744**. A desired output **748** can be taken from a second input **747** to amplifier **744**. The output from amplifier **744** is generally provided at node A, which in turn is connected to the input to transistor **720** and the input to transistor **724**. Note that transistor **724** is generally coupled between ground **701** and first input **733** of amplifier **728**. The second input **731** of amplifier **728** can produce a threshold voltage 760.

Layer **708** provides an output **717** that can be connected to resistor **716**, transistor **725** and a first input **737** to an amplifier **732**. Resistor **716** is generally coupled between ground **701** and the output **717** of layer **708**. The first input **737** of amplifier **732** is also electrically connected to the output **717** of layer **708**. A second input **735** to amplifier **732** may be tied to a threshold voltage 758. The output from amplifier **732** can in turn be fed to an inverting amplifier **738**. The output from inverting amplifier **738** may in turn be provided to a NOR device **742**. Similarly, an output **718** from layer **710** can be connected to a resistor **719**, a transistor **728** and a first input **739** of an amplifier **734**. Note that resistor **719** is generally coupled between node **701** and the output **719** of layer **710**. A second input **741** of amplifier **734** may be coupled to a threshold voltage 762. The output from of NOR device **742** is generally connected to a first input **749** of an amplifier **746**. A desired output **752** can be taken from a second input **751** of amplifier **746**. An actual output **754** can be taken from first input **749** of amplifier **746**. The output of amplifier **746** may be provided at node B, which in turn can be tied back to the respective inputs to transistors **725** and **728**. Note that transistor **725** is generally coupled between ground **701** and the first input **737** of amplifier **732**. Similarly, transistor **728** is generally connected between ground **701** and the first input **739** of amplifier **734**.

Note that transistors **720**, **724**, **725** and/or **728** each can essentially function as a switch to ground. A transistor such as, for example, transistor **720**, **724**, **725** and/or **728** may comprise a field-effect transistor (FET) or another type of transistor, such as, for example, a single-electron transistor (SET). Single-electron transistor (SET) circuits are essential for hybrid circuits combining quantum SET devices with conventional electronic devices. Thus, SET devices and circuits may be adapted for use with the physical neural network of the present invention. This is particularly important because as circuit design rules begin to move into regions of the sub-**100** nanometer scale, where circuit paths are only 0.001 of the thickness of a human hair, prior art device technologies will begin to fail, and current leakage in traditional transistors will become a problem. SET offers a solution at the quantum level, through the precise control of a small number of individual electrons.

Transistors such as transistors **720**, **724**, **725** and/or **728** can also be implemented as carbon nanotube transistors. An example of a carbon nanotube transistor is disclosed in U.S. Patent Application No. 2001/0023986A1 to Macevski, which is dated Sep. 27, 2001 and is entitled, "System and Method for Fabricating Logic Devices Comprising Carbon Nanotube Transistors." U.S. Patent Application No. 2001/0023986A1 does not teach or claim a physical neural network, but instead teaches the formation of a discrete carbon nanotube transistor. Thus, U.S. Patent Application No. 2001/0023986A1 is not considered a limiting feature of the present invention, nor does this reference teach, anticipate or suggest the invention described herein. Instead, this reference is discussed briefly herein for background purposes only and to generally illustrate the use of a particular type of discrete transistor in the nanodomain.

A truth table for the output of circuit **700** is illustrated at block **780** in FIG. **7**. As indicated at block **780**, when an excitatory output is high and the inhibitory output is also high, the final output is low. When the excitatory output is high and the inhibitory output is low, the final output is high. Similarly, when the excitatory output is low and the inhibitory output is high, the final output is low. When the excitatory output is low and the inhibitory output is also low, the final output is low. Note that layers **704** and **708** may thus comprise excitatory connections, while layers **706** and **710** may comprise inhibitory connections.

For every desired output, two sets of connections are used. The output of a two-diode neuron can be fed into an op-amp (e.g., a comparator). If the output that the op-amp receives is low when it should be high, the op-amp outputs a low signal. This low signal can cause the transistors (e.g., transistors **720**, **725**) to saturate and ground out the pre-diode junction for the excitatory diode. This causes, like before, an increase in the voltage drop across those connections that need to increase their strength. Note that only those connections going to the excitatory diode are strengthened. Likewise, if the desired output were low when the actual output was high, the op-amp can output a high signal. This can cause the inhibitory transistor (e.g., an NPN transistor) to saturate and ground out the neuron junction of the inhibitory connections. Connections going to the inhibitory diode can thereafter strengthen.

At all times during the learning process, a weak alternating electric field can be applied perpendicular to the connections. This can cause the connections to weaken by rotating the nanotube perpendicular to the connection direction. This perpendicular field is important because it can allow for a much higher degree of adaptation. To understand this, one must realize that the connections cannot (practically) keep getting stronger and stronger. By weakening those connections not contributing much to the desired output, we decrease the necessary strength of the needed connections and allow for more flexibility in continuous training. This perpendicular alternating voltage can be realized by the addition of two electrodes on the outer extremity of the connection set, such as plates sandwiching the connections (i.e., above and below). Other mechanisms, such as increasing the temperature of the nanotube suspension could also be used for such a purpose, although this method is perhaps a little less controllable or practical.

The circuit depicted in FIG. **7** can be separated into two separate circuits. The first part of the circuit can be composed of nanotube connections, while the second part of the circuit comprises the "neurons" and the learning mechanism (i.e., op-amps/comparator). The learning mechanism on first glance appears similar to a relatively standard circuit that could be implemented on silicon with current technology. Such a silicon implementation can thus comprise the "neuron" chip. The second part of the circuit (i.e., the connections) is thus a new type of chip, although it could be constructed with current technology. The connection chip can be composed of an orderly array of electrodes spaced anywhere from, for example, 100 nm to 1 μm or perhaps even further. In a biological system, one talks of synapses

connecting neurons. It is in the synapses where the information is processed, (i.e., the "connection weights"). Similarly, such a chip can contain all of the synapses for the physical neural network. A possible arrangement thereof can be seen in FIG. **8**.

FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., connection network **800**) that can be implemented in accordance with the present invention. FIG. **8** thus illustrates a possible chip layout for a connection chip (i.e., s connection network **800** that can be implemented in accordance with the present invention. Chip layout **800** can include an input array composed of a plurality of inputs **801**, **802**, **803**, **804**, and **805**, which are generally provided to a plurality of layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. A plurality of outputs **802** can be derived from layers **806**, **807**, **808**, **809**, **810**, **811**, **812**, **813**, **814**, and **815**. Inputs **801** can be coupled to layers **806** and **807**, while inputs **802** can be connected to layers **808** and **809**. Similarly, inputs **803** can be connected to layers **810** and **811**. Also, inputs **804** are generally connected to layers **812** and **813**. Inputs **805** are generally connected to layers **814** and **815**.

Similarly, such an input array can includes a plurality of inputs **831**, **832**, **833**, **834** and **835** which are respectively input to a plurality of layers **816**, **817**, **818**, **819**, **820**, **821**, **822**, **823**, **824** and **825**. Thus, inputs **831** are connected to layers **816** and **817**, while inputs **832** are coupled to layers **818** and **819**. Additionally, inputs **833** are connected to layers **820** and **821**. Inputs **834** are connected to layers **822** and **823**. Finally, inputs **835** are connected to layers **824** and **825**. Arrows **828** and **830** represent a continuation of the aforementioned connection network pattern. Those skilled in the art can appreciate, of course, that chip layout **800** is not intended to represent an exhaustive chip layout or to limit the scope of the invention. Many modifications and variations to chip layout **800** are possible in light of the teachings herein without departing from the scope of the present invention. It is contemplated that the use of a chip layout, such as chip layout **800**, can involve a variety of components having different characteristics.

Preliminary calculations by the present inventor based on a maximum etching capability of 200 nm resolution have indicated that over 4 million synapses could fit on an area of approximately 1 cm^2. The smallest width that an electrode can possess is generally based on current lithography. Such a width may of course change as the lithographic arts advance. This value is actually about 70 nm for state-of-the-art techniques currently. These calculations are of course extremely conservative, and are not considered a limiting feature of the present invention. Such calculations are based on an electrode with, separation, and gap of approximately 200 nm. For such a calculation, 166 connection networks comprising 250 inputs and 100 outputs can fit within a one square centimeter area.

If such chips are stacked vertically, an untold number of synapses could be attained. This is two to three orders of magnitude greater than some of the most capable neural network chips out there today, chips that rely on standard methods to calculate synapse weights. Of course, the geometry of the chip could take on many different forms, and it is quite possible (i.e., based on a conservative lithography and chip layout) that many more synapses could fit within the same space. The training of a neural network chip of this size would take a fraction of the time of a comparably sized traditional chip using digital technology.

The training of such a chip is primarily based on two assumptions. First, the inherent parallelism of a physical neural network (i.e., a Knowm) can permit all training sessions to occur simultaneously, no matter now large the associated connection network. Second, recent research has indicated that near perfect aligning of nanotubes can be accomplished, for example, in approximately 15 minutes. If one considers that the input data, arranged as a vector of binary "highs and low's" is presented to the Knowm simultaneously, and that all training vectors are presented one after the other in rapid succession (e.g., 100 MHz or more), then each connection would "see" a different frequency in direct proportion to the amount of time that its connection is required for accurate data processing (i.e., provided by a feedback mechanism). Thus, if it only takes for example, approximately 15 minutes to attain an almost perfect state of alignment, then this amount of time would comprise the longest amount of time required to train, assuming that all of the training vectors are presented during that particular time period.

FIG. **9** illustrates a flow chart **900** of operations illustrating operational steps that may be followed to construct a connection network, in accordance with a preferred embodiment of the present invention. Initially, as indicated at block **902**, a connection gap is created from a connection network structures. As indicated earlier, the goal for such a connection network is generally to develop a network of connections of "just" the right values to satisfy particular information processing requirements, which is precisely what a neural network accomplishes. As illustrated at block **904**, a solution is prepared, which is composed of nanoconductors and a "solvent." Note that the term "solvent" as utilized herein has a variable meaning, which includes the traditional meaning of a "solvent," and also a suspension.

The solvent utilized can comprise a volatile liquid that can be confined or sealed and not exposed to air. For example, the solvent and the nanoconductors present within the resulting solution may be sandwiched between wafers of silicon or other materials. If the fluid has a melting point that is approximately at room temperature, then the viscosity of the fluid could be controlled easily. Thus, if it is desired to lock the connection values into a particular state, the associated physical neural network (i.e., Knowm) may be cooled slightly until the fluid freezes. The term "solvent" as utilized herein thus can include fluids such as for example, toluene, hexadecane, mineral oil, etc. Note that the solution in which the nanoconductors (i.e., nanoconnections) are present should generally comprise a dielectric. Thus, when the resistance between the electrodes is measured, the conductivity of the nanoconductors is essentially measured, not that of the solvent. The nanoconductors can be suspended in the solution or can alternately lie on the bottom surface of the connection gap. The solvent may also be provided in the form of a gas.

As illustrated thereafter at block **906**, the nanoconductors must be suspended in the solvent, either dissolved or in a suspension of sorts, but generally free to move around, either in the solution or on the bottom surface of the gap. As depicted next at block **908**, the electrical conductance of the solution must be less than the electrical conductance of the suspended nanoconductor(s). Similarly, the electrical resistance of the solution is greater than the electrical resistance of the nanoconductor.

Next, as illustrated at block **910**, the viscosity of the substance should not be too much so that the nanoconductors cannot move when an electric field (e.g., voltage) is applied. Finally, as depicted at block **912**, the resulting solution of the "solvent" and the nanoconductors is thus located within the connection gap.

Note that although a logical series of steps is illustrated in FIG. **9**, it can be appreciated that the particular flow of steps can be re-arranged. Thus, for example, the creation of the connection gap, as illustrated at block **902**, may occur after the preparation of the solution of the solvent and nanoconductor(s), as indicated at block **904**. FIG. **9** thus represents merely possible series of steps, which may be followed to create a connection network. It is anticipated that a variety of other steps may be followed as long as the goal of achieving a connection network in accordance with the present invention is achieved. Similar reasoning also applies to FIG. **10**.

FIG. **10** depicts a flow chart **1000** of operations illustrating operational steps that may be utilized to strengthen nanoconductors within a connection gap, in accordance with a preferred embodiment of the present invention. As indicated at block **1002**, an electric field can be applied across the connection gap discussed above with respect to FIG. **9**. The connection gap can be occupied by the solution discussed above. As indicated thereafter at block **1004**, to create the connection network, the input terminals can be selectively raised to a positive voltage while the output terminals are selectively grounded. As illustrated thereafter at block **1006**, connections thus form between the inputs and the outputs. The important requirements that make the resulting physical neural network functional as a neural network is that the longer this electric field is applied across the connection gap, or the greater the frequency or amplitude, the more nanoconductors align and the stronger the connection becomes. Thus, the connections that get utilized the most frequently become the strongest.

As indicated at block **1008**, the connections can either be initially formed and have random resistances or no connections will be formed at all. By forming initial random connections, it might be possible to teach the desired relationships faster, because the base connections do not have to be built up as much. Depending on the rate of connection decay, having initial random connections could prove to be a faster method, although not necessarily. A connection network will adapt itself to whatever is required regardless of the initial state of the connections. Thus, as indicated at block **1010**, as the electric field is applied across the connection gap, the more the nonconductor(s) will align and the stronger the connection becomes. Connections (i.e., synapses) that are not used are dissolved back into the solution, as illustrated at block **1012**. As illustrated at block **1014**, the resistance of the connection can be maintained or lowered by selective activations of the connections. In other words, "if you do not use the connection, it will fade away," much like the connections between neurons in a human brain.

The neurons in a human brain, although seemingly simple when viewed individually, interact in a complicated network that computes with both space and time. The most basic picture of a neuron, which is usually implemented in technology, is a summing device that adds up a signal. Actually, this statement can be made even more general by stating that a neuron adds up a signal in discrete units of time. In other words, every group of signals incident upon the neuron can be viewed as occurring in one moment in time. Summation thus occurs in a spatial manner. The only difference between one signal and another signal depends on where such signals originate. Unfortunately, this type of data processing excludes a large range of dynamic, varying situations that cannot necessarily be broken up into discrete units of time.

The example of speech recognition is a case in point. Speech occurs in the time domain. A word is understood as the temporal pronunciation of various syllables. A sentence is composed of the temporal separation of varying words. Thoughts are composed of the temporal separation of varying sentences. Thus, for an individual to understand a spoken language at all, a syllable, word, sentence or thought must exert some type of influence on another syllable, word, sentence or thought. The most natural way that one sentence can exert any influence on another sentence, in the light of neural networks, is by a form of temporal summation. That is, a neuron "remembers" the signals it received in the past.

The human brain accomplishes this feat in an almost trivial manner. When a signal reaches a neuron, the neuron has an influx of ions rush through its membrane. The influx of ions contributes to an overall increase in the electrical potential of the neuron. Activation is achieved when the potential inside the cell reaches a certain threshold. The one caveat is that it takes time for the cell to pump out the ions, something that it does at a more or less constant rate. So, if another signal arrives before the neuron has time to pump out all of the ions, the second signal will add with the remnants of the first signal and achieve a raised potential greater than that which could have occurred with only the second signal. The first signal influences the second signal, which results in temporal summation.

Implementing this in a technological manner has proved difficult in the past. Any simulation would have to include a "memory" for the neuron. In a digital representation, this requires data to be stored for every neuron, and this memory would have to be accessed continually. In a computer simulation, one must discritize the incoming data, since operations (such as summations and learning) occur serially. That is, a computer can only do one thing at a time. Transformations of a signal from the time domain into the spatial domain require that time be broken up into discrete lengths, something that is not necessarily possible with real-time analog signals in which no point exists within a time-varying signal that is uninfluenced by another point.

A physical neural network, however, is generally not digital. A physical neural network is a massively parallel analog device. The fact that actual molecules (e.g., nanoconductors) must move around (in time) makes temporal summation a natural occurrence. This temporal summation is built into the nanoconnections. The easiest way to understand this is to view the multiplicity of nanoconnections as one connection with one input into a neuron-like node (Op-amp, Comparator, etc.). This can be seen in FIG. **11**.

FIG. **11** illustrates a schematic diagram of a circuit **1100** illustrating temporal summation within a neuron, in accordance with a preferred embodiment of the present invention. As indicated in FIG. **11**, an input **1102** can be provided to a nanoconnections **1104**, which in turn can provide a signal, which can be input to an amplifier **1110** (e.g., op amp) at node B. A resistor **1106** is connected to node A, which in turn is electrically equivalent to node B. Node B is connected to a negative input of amplifier **1100**. Resistor **1108** is also connected to a ground **1108**. Amplifier **1110** provides output **1114**. Note that although nanoconnections **1104** is referred to in the plural it can be appreciated that nanoconnections **1104** can comprise a single nanoconnection or a plurality of nanoconnections. For simplicity sake, however, the plural form is used to refer to nanoconnections **1104**.

Input **1102** can be provided by another physical neural network (i.e., Knowm) to cause increased connection strength of nanoconnections **1104** over time. This input would most likely arrive in pulses, but could also be

continuous. A constant or pulsed electric field perpendicular to the connections would serve to constantly erode the connections, so that only signals of a desired length or amplitude could cause a connection to form. Once the connection is formed, the voltage divider formed by nanoconnection **1104** and resistor **1106** can cause a voltage at node A in direct proportion to the strength of nanoconnections **1104**. When the voltage at node A reaches a desired threshold, the amplifier (i.e., an op-amp and/or comparator), will output a high voltage (i.e., output **1114**). The key to the temporal summation is that, just like a real neuron, it takes time for the electric field to breakdown the nanoconnections **1104**, so that signals arriving close in time will contribute to the firing of the neuron (i.e., op-amp, comparator, etc.). Temporal summation has thus been achieved. The parameters of the temporal summation could be adjusted by the amplitude and frequency of the input signals and the perpendicular electric field.

FIG. **12** depicts a block diagram illustrating a pattern recognition system **1200**, which may be implemented with a physical neural network device **1222**, in accordance with an alternative embodiment of the present invention. Note that pattern recognition system **1200** can be implemented as a speech recognition system. Those skilled in the art can appreciate, however, that although pattern recognition system **1200** is depicted herein in the context of speech recognition, a physical neural network device (i.e., a Knowm device) may be implemented with other pattern recognition systems, such as visual and/or imaging recognition systems. FIG. **12** thus does not comprise a limiting feature of the present invention and is presented for general edification and illustrative purposes only. Those skilled in the art can appreciate that the diagram depicted in FIG. **12** may be modified as new applications and hardware are developed. The development or use of a pattern recognition system such as pattern recognition system **1200** of FIG. **12** by no means limits the scope of the physical neural network (i.e., Knowm) disclosed herein.

FIG. **12** thus illustrates in block diagram fashion, the system structure of a speech recognition device using a neural network according to an alternative embodiment of the present invention. The pattern recognition system **1200** is provided with a CPU **1211** for performing the functions of inputting vector rows and instructor signals (vector rows) to an output layer for the learning process of a physical neural network device **1222**, and changing connection weights between respective neuron devices based on the learning process. Pattern recognition system **1200** can be implemented within the context of a data-processing system, such as, for example, a personal computer or personal digital assistant (PDA), both of which are well known in the art.

The CPU **1211** can perform various processing and controlling functions, such as pattern recognition, including but not limited to speech and/or visual recognition based on the output signals from the physical neural network device **1222**. The CPU **1211** is connected to a read-only memory (ROM) **1213**, a random-access memory (RAM) **1214**, a communication control unit **1215**, a printer **1216**, a display unit **1217**, a keyboard **1218**, an FFT (fast Fourier transform) unit **1221**, a physical neural network device **1222** and a graphic reading unit **1224** through a bus line **1220** such as a data bus line. The bus line **1220** may comprise, for example, an ISA, EISA, or PCI bus.

The ROM **1213** is a read-only memory storing various programs or data used by the CPU **1211** for performing processing or controlling the learning process, and speech recognition of the physical neural network device **1222**. The ROM **1213** may store programs for carrying out the learning process according to error back-propagation for the physical neural network device or code rows concerning, for example, 80 kinds of phonemes for performing speech recognition. The code rows concerning the phonemes can be utilized as second instructor signals and for recognizing phonemes from output signals of the neuron device network. Also, the ROM **1213** can store programs of a transformation system for recognizing speech from recognized phonemes and transforming the recognized speech into a writing (i.e., written form) represented by characters.

A predetermined program stored in the ROM **1213** can be downloaded and stored in the RAM **1214**. RAM **1214** generally functions as a random access memory used as a working memory of the CPU **1211**. In the RAM **1214**, a vector row storing area can be provided for temporarily storing a power obtained at each point in time for each frequency of the speech signal analyzed by the FFT unit **1221**. A value of the power for each frequency serves as a vector row input to a first input portion of the physical neural network device **1222**. Further, in the case where characters or graphics are recognized in the physical neural network device, the image data read by the graphic reading unit **1224** are stored in the RAM **1214**.

The communication control unit **1215** transmits and/or receives various data such as recognized speech data to and/or from another communication control unit through a communication network **1202** such as a telephone line network, an ISDN line, a LAN, or a personal computer communication network. Network **1202** may also comprise, for example, a telecommunications network, such as a wireless communications network. Communication hardware methods and systems thereof are well known in the art.

The printer **1216** can be provided with a laser printer, a bubble-type printer, a dot matrix printer, or the like, and prints contents of input data or the recognized speech. The display unit **1217** includes an image display portion such as a CRT display or a liquid crystal display, and a display control portion. The display unit **1217** can display the contents of the input data or the recognized speech as well as a direction of an operation required for speech recognition utilizing a graphical user interface (GUI).

The keyboard **1218** generally functions as an input unit for varying operating parameters or inputting setting conditions of the FFT unit **1221**, or for inputting sentences. The keyboard **1218** is generally provided with a ten-key numeric pad for inputting numerical figures, character keys for inputting characters, and function keys for performing various functions. A mouse **1219** can be connected to the keyboard **1218** and serves as a pointing device.

A speech input unit **1223**, such as a microphone can be connected to the FFT unit **1221**. The FFT unit **1221** transforms analog speech data input from the voice input unit **1223** into digital data and carries out spectral analysis of the digital data by discrete Fourier transformation. By performing a spectral analysis using the FFT unit **1221**, the vector row based on the powers of the respective frequencies are output at predetermined intervals of time. The FFT unit **1221** performs an analysis of time-series vector rows, which represent characteristics of the inputted speech. The vector rows output by the EFT **1221** are stored in the vector row storing area in the RAM **1214**. The graphic reading unit **224**, provided with devices such as a CCD (Charged Coupled Device), can be used for reading images such as characters or graphics recorded on paper or the like. The image data read by the image-reading unit **1224** are stored in the RAM

1214. Note that an example of a pattern recognition apparatus, which may be modified for use with the physical neural network of the present invention, is disclosed in U.S. Pat. No. 6,026,358 to Tomabechi, Feb. 16, 2000, "Neural Network, A Method of Learning of a Neural Network and Phoneme Recognition Apparatus Utilizing a Neural Network." It can be appreciated by those skilled in the art that the Tomabechi reference does not teach, suggest or anticipate the invention disclosed herein. The Tomabechi reference is discussed herein for illustrative, background general edification purposes only and is not considered a limiting feature of the present invention.

The implications of a physical neural network are tremendous. With existing lithography technology, many electrodes in an array such as depicted in FIG. 5 can be etched onto a wafer of silicon. The neurons (i.e., op-amps, diodes, etc.), as well as the training circuitry illustrated in FIG. 6, could be built onto the same silicon wafer, although it may be desirable to have the connections on a separate chip due to the liquid solution of nanoconductors. A solution of suspended nanoconductors could be placed between the electrode connections and the chip could be packaged. The resulting "chip" would look much like a current Integrated Chip (IC) or VLSI (very large scale integrated) chips. One could also place a rather large network parallel with a computer processor as part of a larger system. Such a network, or group of networks, could add significant computational capabilities to standard computers and associated interfaces.

For example, such a chip may be constructed utilizing a standard computer processor in parallel with a large physical neural network or group of physical neural networks. A program can then be written such that the standard computer teaches the neural network to read, or create an association between words, which is precisely the same sort of task in which neural networks can be implemented. Once the physical neural network is able to read, it can be taught for example to "surf" the Internet and find material of any particular nature. A search engine can then be developed that does not search the Internet by "keywords", but instead by meaning. This idea of an intelligent search engine has already been proposed for standard neural networks, but until now has been impractical because the network required was too big for a standard computer to simulate. The use of a physical neural network (i.e., physical neural network) as disclosed herein now makes a truly intelligent search engine possible.

A physical neural network can be utilized in other applications, such as, for example, speech recognition and synthesis, visual and image identification, management of distributed systems, self-driving cars and filtering. Such applications have to some extent already been accomplished with standard neural networks, but are generally limited in expense, practicality and not very adaptable once implemented. The use of a physical neural network can permit such applications to become more powerful and adaptable. Indeed, anything that requires a bit more "intelligence" could incorporate a physical neural network. One of the primary advantages of a physical neural network is that such a device and applications thereof can be very inexpensive to manufacture, even with present technology. The lithographic techniques required for fabricating the electrodes and channels therebetween has already been perfected and implemented in industry.

Most problems in which a neural network solution is implemented are complex adaptive problems, which change in time. An example is weather prediction. The usefulness of a physical neural network is that it could handle the enormous network needed for such computations and adapt itself in real-time. An example wherein a physical neural network (i.e., Knowm) can be particularly useful is the Personal Digital Assistant (PDA). PDA's are well known in the art. A physical neural network applied to a PDA device can be advantageous because the physical neural network can ideally function with a large network that could constantly adapt itself to the individual user without devouring too much computational time from the PDA. A physical neural network could also be implemented in many industrial applications, such as developing a real-time systems control to the manufacture of various components. This systems control can be adaptable and totally tailored to the particular application, as necessarily it must.

The embodiments and examples set forth herein are presented to best explain the present invention and its practical application and to thereby enable those skilled in the art to make and utilize the invention. Those skilled in the art, however, will recognize that the foregoing description and examples have been presented for the purpose of illustration and example only. Other variations and modifications of the present invention will be apparent to those of skill in the art, and it is the intent of the appended claims that such variations and modifications be covered. The description as set forth is not intended to be exhaustive or to limit the scope of the invention. Many modifications and variations are possible in light of the above teaching without departing from the scope of the following claims. It is contemplated that the use of the present invention can involve components having different characteristics. It is intended that the scope of the present invention be defined by the claims appended hereto, giving full cognizance to equivalents in all respects.

The embodiments of an invention in which an exclusive property or right is claimed are defined as follows:

1. A physical neural network based on nanotechnology, said physical neural network comprising:

at least one neuron-like node that sums at least one input signal and generates at least one output signal based on a threshold associated with said at least one input signal; and

at least one connection network associated with said at least one neuron-like node wherein said at least one connection network comprises a plurality of interconnected nanoconnections, such that each nanoconnection of said plurality of interconnected nanoconnections is strengthened or weakened according to an application of an electric field.

2. The physical neural network of claim 1 wherein said at least one output signal comprises a non-linear output signal based on said threshold.

3. The physical network of claim 1 wherein said at least one output signal comprises a linear output signal based on said threshold.

4. The physical neural network of claim 1 wherein said threshold comprises a threshold below which said at least one output signal is not generated and above which said at least one output signal is generated.

5. The physical neural network of claim 1 wherein said at least one connection network comprises:

a number of layers of said nanoconnections;

wherein said number of layers is equal to a number of desired outputs from said at least one connection network; and

wherein said nanoconnections are formed without influence by disturbances resulting from other nanoconnections thereof.

6. The physical neural network of claim 1 wherein at least one nanoconnection of said plurality of interconnected nanoconnections comprises an electrically conducting material.

7. The physical neural network of claim 6 wherein electrically conducting material is chosen such that a dipole is induced in said electrically conducting material in the presence of an electric field.

8. The physical neural network of claim 6 wherein said electrically conducting material comprises a chemically induced permanent dipole.

9. The physical neural network of claim 1 wherein said at least one nanoconnection comprises at least one nanoconductor.

10. The physical neural network of claim 9 wherein said at least one connection network comprises:

at least one connection network structure having a connection gap formed therein; a solution located within said connection gap;

wherein said solution comprises a solvent and said at least one nanoconductor; and

wherein an electric field applied across said connection gap to permit an alignment of at least one nanoconductor within said connection gap.

11. The physical neural network of claim 10 wherein said at is at least one nanoconductor is suspended in said solvent.

12. The physical neural network of claim 10 wherein said at least one nanoconductor is located at a bottom of said connection gap.

13. The physical neural network of claim 10 wherein an electrical conductance of said solution is less than an electrical conductance of said at least one nanoconductor within said solution.

14. The physical neural network of claim 10 wherein a viscosity of said solution permits said at least one nanoconductor to move when said electric field is applied across said connection gap.

15. The physical neural network of claim 10 wherein said at least one nanoconductor experiences an increased alignment in accordance with an increase in said electric field applied across said connection gap.

16. The physical neural network of claim 15 wherein nanoconnections of said at least one neuron-like node that are utilized most frequently by said at least one neuron-like node become stronger with each use thereof.

17. The physical neural network of claim 16 wherein said nanoconnections that are utilized least frequently become increasingly weak and eventually become unaligned.

18. The physical neural network of claim 16 wherein said at least one nanoconnection comprises a resistance, which is raised or lowered by a selective activation of said at least one nanoconnection.

19. The physical neural network of claim 9 wherein said at least one nanoconductor comprises a nanowire.

20. The physical neural network of claim 9 wherein said at least one nanoconductor comprises a nanotube.

21. The physical neural network of claim 9 wherein said at least one nanoconductor comprises a plurality of nanoparticles.

22. The physical neural network of claim 1 wherein at least one nanoconnection of said at least one connection network comprises a negative connection associated with said at least one neuron-like node.

23. The physical neural network of claim 22 wherein said at least one connection network comprises:

a number of layers of nanoconnections;

wherein said number of layers is equal to a number of desired outputs from said at least one connection network;

wherein said number of layers is equal to twice a number of desired outputs form said at least one connection network, if said nanoconnections comprise negative connections thereof; and

wherein said nanoconnections are formed without influence by disturbances resulting from other nanoconnections thereof.

24. A physical neural network apparatus based on nanotechnology, said physical neural apparatus comprising:

at least one connection network associated with at least one neuron-like node wherein said at least one connection network comprises a plurality of interconnected nanoconductors, such that each nanoconductor of said plurality of interconnected nanoconductors is strengthened or weakened according to an application of an electric field;

wherein each nanoconductor of said plurality of interconnected nanoconductors experiences an increase in alignment in accordance with an increase in said electric field;

wherein nanoconductors of said plurality of interconnected nanoconductors that are utilized most frequently by said at least one neuron-like node became stronger with each use thereof; and

wherein nanoconductors of said plurality of interconnected nanoconductors that are utilized least frequently become increasingly weak and eventually become unaligned.

25. The physical neural network apparatus of claim 24 wherein said at least one neuron-like node sums at least one input signal and generates at least one output signal based on a threshold associated with said at least one input signal.

26. The physical neural network of claim 24 wherein said at least one connection network comprises a plurality of layers, wherein a number of layers of said plurality of layers is equal to a desired number of outputs from said at least one connection network.

27. A method for assembling a physical neural network based on nanotechnology, said method comprising the steps of:

forming at least one neuron-like node that sums at least one input signal and generates at least one output signal based on a threshold associated with said at least one input signal; and

configuring at least one connection network associated with said at least one neuron-like node wherein said at least one connection network comprises a plurality of interconnected nanoconnections, such that each nanoconnection of said plurality of interconnected nanoconnections is strengthened or weakened according to an application of an electric field.

28. The method of claim 27 wherein said at least one output signal comprises a non-linear output signal based on said threshold.

29. The method of claim 27 wherein said at least one output signal comprises a linear output signal based on said threshold.

30. The method of claim 27 wherein said threshold comprises a threshold below which said at least one output signal is not generated and above which said at least one output signal is generated.

31. The method of claim 27 further comprising the steps of:

configuring said at least one connection network to comprise a number of layers of said nanoconnections, wherein said number of layers is equal to a number of desired outputs from said at least one connection network; and

forming said nanoconnections without influence by disturbances resulting from other nanoconnections thereof.

32. The method of claim **27** comprising the step of; forming at least one nanoconnection of said plurality of interconnected nanoconnections from an electrically conducting material.

33. The method of claim **32** further comprising the step of: configuring said electrically conducting material such that a dipole is induced in said electrically conducting material in the presence of an electric field.

34. The method of claim **32** comprising the step of: chemically inducing a permanent dipole within said electrically conducting material.

35. The method of claim **27** wherein said at least one nanoconnection comprises at least one nanoconductor.

36. The method of claim **35** further comprising the steps of:

forming a connection gap from at least one connection network structure associated with said at least one connection network;

locating a solution within said connection gap; configuring said solution to comprises a solvent and said at least one nanoconductor; and

applying an electric field across said connection gap to permit an alignment of at least one nanoconductor within said connection gap.

37. The method of claim **36** further comprising the step of: suspending said at least one nanoconductor in said solvent.

38. The method of claim **36** further comprising the step of: locating said at least one nanoconductor at a bottom of said connection gap.

39. The method of claim **36** further comprising the step of: configuring said solution such that an electrical conductance of said solution is less than an electrical conductance of said at least one nanoconductor within said solution.

40. The method of claim **36** further comprising the step of: configuring said solution to comprise a viscosity that permits said at least one nanoconductor to move when said electric field is applied across said connection gap.

41. The method of claim **36** further comprising the step of: increasing an application of said electric field across said connection gap to thereby permit said at least one nanoconductor to experience an increased alignment in accordance with an increase in said electric field applied across said connection gap.

42. The method of claim **41** wherein nanoconnections of said at least one neuron-like node that are utilized most frequently by said at least one neuron-like node become stronger with each use thereof.

43. The method of claim **42** wherein said nanoconnections that are utilized least frequently become increasingly weak and eventually become unaligned.

44. The method of claim **42** wherein said at least one nanoconnection comprises a resistance, which is raised or lowered by a selective activation of said at least one nanoconnection.

45. The method of claim **35** further comprising the step of: configuring said at least one nanoconductor to comprise a nanowire.

46. The method of claim **35** further comprising the step of: configuring said at least one nanoconductor to comprise a nanotube.

47. The method of claim **35** further comprising the step of: configuring said at least one nanoconductor to comprise a plurality of nanoparticles.

48. The method of claim **27** wherein at least one nanoconnection of said at least one connection network comprises a negative connection associated with said at least one neuron-like node.

49. The method of claim **48** further comprising the step of:

configuring said at least one connection network to comprise a number of layers of nanoconnections, wherein said number of layers is equal to a number of desired outputs from said at least one connection network;

wherein said number of layers is equal to twice a number of desired outputs form said at least one connection network, if said nanoconnections comprise negative connections thereof; and wherein said nanoconnections are formed without influence by disturbances resulting from other nanoconnections thereof.

50. A method for assembling a physical neural network apparatus based on nanotechnology, said method comprising the steps of:

forming at least one connection network associated with at least one neuron-like node wherein said at least one connection network comprises a plurality of interconnected nanoconductors, such that each nanoconductor of said plurality of interconnected nanoconductors is strengthened or weakened according to an application of an electric field;

wherein each nanoconductor of said plurality of interconnected nanoconductors experiences an increase in alignment in accordance with an increase in said electric field; wherein nanoconductors of said plurality of interconnected nanoconductors that are utilized most frequently by said at least one neuron-like node become stronger with each use thereof; and

wherein nanoconductors of said plurality of interconnected nanoconductors that are utilized least frequently become increasingly weak and eventually become unaligned.

51. The method apparatus of claim **50** further comprising the steps of: summing at least one input signal via said at least one neuron-like node; and generating at least one output signal based on a threshold associated with said at least one input signal.

52. The method of claim **50** further comprising the step of: configuring said at least one connection network to comprise a plurality of layers, wherein a number of layers of said plurality of layers is equal to a desired number of outputs from said at least one connection network.

* * * * *

In-situ and real time determination of metallic and semi-conducting SWNTs in suspension via dielectrophoresis

Natacha Mureau[1], *Ernest Mendoza*[1], *Kai F. Hoettges*[2], *S.R.P. Silva*[1], *Michael P. Hughes*[2*]

[1]University of Surrey, School of Electronics and Physical Sciences, Nano-electronics centre, Advanced Technology Institute, Guildford, Surrey GU2 7XH, England

[2]University of Surrey, School of Engineering, Centre for Biomedical Engineering, Guildford, Surrey GU2 7XH, England

* Corresponding author. Email: M.Hughes@surrey.ac.uk

RECEIVED DATE (to be automatically inserted after your manuscript is accepted if required according to the journal that you are submitting your paper to)

Abstract

In this letter we report a new method combining dielectrophoresis and impedance spectroscopy to provide rapid, accurate measurement of dielectrophoretic collection of SWNTs in real time. We used a Triton X-100 suspension of mixed SWNTs to demonstrate their precise dielectrophoretic collection behavior. Results indicate our sample contains 21.5 % of metallic and 78.5 % of semi-conducting carbon nanotubes, and that the range 1 MHz-15 MHz is ideal to collect only the metallic ones. We used optical absorption to confirm these proportions.

Since their discovery by Iijima in 1991[1], carbon nanotubes have demonstrated great potential in nanoelectronic applications due to their remarkable chemical, physical and electrical properties[2-4]. In particular, depending on their diameter and chirality, single-wall carbon nanotubes (SWNTs) can behave either as a metal or a semi-conductor[5]. Dielectrophoresis[6-7] (DEP), the movement of particles in non-uniform electric fields, has been used to separate mixtures of semi-conducting and metallic SWNTs[8-10]. Measurement of the frequency-dependent collection of dielectrophoresis has thus far been based on Raman spectroscopy, which offers low frequency resolution and requires some complex interpretation of results[11-14]. In this paper we report the application of combined dielectrophoresis and impedance measurements to provide real-time rapid, accurate measurement of dielectrophoretic collection of SWNTs. Using mixed SWNTs suspended in a Triton X-100 solution, we have demonstrated their precise dielectrophoretic collection behavior, which can be used to provide an accurate assessment of the optimum conditions of DEP separation and determination of the dielectric properties of the carbon nanotubes. Results indicate the sample analyzed in this study contains 21.5 % of metallic and 78.5 % of semi-conducting carbon nanotubes before separation, and that the range 1 MHz-15 MHz is optimal to collect only the metallic type of SWNTs. The proportion of the type of SWNTs present in the sample has been confirmed using optical absorption studies.

The measurements were performed on needle-shaped electrodes with an inter-electrode gap of 10 μm and energized with a sinusoidal 10 $V_{pk\text{-}pk}$ voltage, at least 5 frequencies per decade over the range 10 kHz-20 MHz using a TG120 20 MHz signal generator. A resistance of 4.33 kΩ was connected in series with the electrodes, and the voltage measured across the resistor and resistor/electrode combination using an ISO-TECH IDS710 digital oscilloscope. The output of the oscilloscope was input to a computer, where a MATLAB program (The Mathworks, Natick, MA, USA) calculated the impedance, and hence derived the resistance of the inter-electrode gap at intervals of around 1 second. The recording process preceded the SWNT containing suspension being placed on the electrodes, and

continued for typically 5 min. The time constant of the change of resistance as a function of time was then obtained using commercial software. Figure 1 shows a photograph of the electrodes after the application of the nanotubes solution (3 min of exposure to the electric field) clearly showing the concentration of nanotubes between the electrodes, aligned along the electric field. A control experiment with a Triton solution free of carbon nanotubes was also performed to observe a zero collection reference.

Figure 2a shows a representative selection of the impedance versus time data which has been normalized to the electrode impedance. The control solution was weakly conductive (13.7 μS/cm compared with deionized water 7.5 μS/cm) and, after the application of the suspension, a 30 % resistance decrease was observed. Of the nanotube samples tested, two different types of behavior were observed. The first is the group of curves recorded at high frequencies, where after an initial drop at the application of the sample, the values decrease exponentially to a stable value at approximately 200 s. The decrease in resistance (final with respect to initial values) was 53 % for 1 MHz, 39 % for 2 MHz, 48 % for 5 MHz. In a second set of experiments, recorded for low frequencies, the reduction is much more significant, e.g. 88 % for 20 kHz, 79 % for 50 kHz, 67 % for 100 kHz. At 20 MHz, the response indicated that the impedance change was due to the medium alone – that is, the result was identical to the control measurement, indicating no collection of nanotubes is expected.

If we hypothesize that the time taken for the impedance to change due to the collection of nanotubes is inversely proportional to the force acting on those tubes, then the reciprocal of the time constant would indicate the magnitude of the force. A similar approach has been taken using fluorescence measurements of latex beads to determine their dielectric properties[15]. Across a population of particles, this reciprocal also indicates the relative population of particles where multiple populations are present. By modeling the population using known models of dielectrophoretic behavior, it is possible to determine both the electrical properties of each population and the relative number[16].

A plot of the reciprocal collection time constant is shown in figure 2b. As observed, two dielectric dispersions are evident, centered approximately at 250 kHz and 13 MHz. This is consistent with

behavior we would expect for a heterogeneous mix of metallic and semi-conducting SWNTs. We use modeling of the dielectrophoretic response to determine the relative polarizability of the populations, and hence their respective dielectric properties. For a spherical particle of radius r the DEP force is given by equation (1):

$$F_{sphere} = 2\pi r^3 \varepsilon_m Re\left[\frac{\varepsilon_p^* - \varepsilon_m^*}{\varepsilon_p^* + 2\varepsilon_m^*}\right]\nabla E^2 \qquad (1)$$

where, ε_m is the absolute permittivity of the suspending medium, E is the local (rms) electric field, ∇ is the differential vector operator, ε_p^* and ε_m^* are the complex permittivities of the particle and medium respectively, $\varepsilon^* = \varepsilon - j\sigma/\omega$, where, ε is the permittivity, σ the conductivity, ω the angular frequency of the applied field, $j = \sqrt{-1}$, and Re denotes the real part. For the case where the 'spherical' particle is replaced with a rod whose major axis is r_1 and minor axis r_2, the force is given by the equation (2):

$$F_{rod} = \frac{2\pi r_1 r_2^2 \varepsilon_m}{3} Re\left[\frac{\varepsilon_p^* - \varepsilon_m^*}{\varepsilon_m^*}\right]\nabla E^2 \qquad (2).$$

Based on the relative magnitudes of ε_p^* and ε_m^*, which are in turn related to ω, the DEP force acting on a particle can cause it to move either towards or away from high-field regions at electrode edges. These two effects are termed *positive* and *negative* DEP respectively.

The solid line on figure 2b represents the best fit for two populations of nanotubes, superimposed to determine the net frequency-dependent SWNT collection. A spherical model, postulated by Krupke et al.[8] to account for ballistic ion transport, was found to provide a good fit for the conducting nanotubes, when the nanotubes were assigned a conductivity of 130 mSm^{-1} and a relative permittivity of less than $40\varepsilon_0$. We were able to obtain a unique set of parameters that enabled us to model semi-conducting

nanotubes either as spheres or as long, thin ellipsoids – an approach that has been used successfully for dielectrophoretic modeling of rod-shaped nanoparticles in the past[17]. In the case of a spherical model, the nanotubes had a conductivity of 2.2 mSm^{-1}, and for the prolate ellipsoid model, 2.5 mSm^{-1}. As with the conducting nanotubes, the model indicates a permittivity of less than $40\varepsilon_0$. Notably, this permittivity is considerably lower than the near-infinite value suggested in the literature[18]. We postulate that the values are more consistent when considered as an ensemble of nanotubes plus its Debye atmosphere, which dominates the dielectric properties of nanoparticles in aqueous solutions[19-21]. It is also known[21] that the electrical double layer is a low-permittivity environment, the radius of which is of the order of the diameter of the SWNTs themselves. The presence of Triton on the SWNTs may also affect this value. Since the nanotubes vary only in chirality, we would expect both metallic and semi-conducting SWNTs to have identical electrical Debye layer properties.

Once the electrical properties of the SWNTs populations have been obtained, we can further determine the relative populations of the particles by determining the relative multiples of each required to match the spectra in figure 2b. Our calculations indicate that where the semi-conducting nanotubes are modeled as prolate ellipsoids, the proportions of the two polarizability plots in figure 2b indicate the population contains (78.5 ± 1 %) semi-conducting SWNTs, with the remainder being metallic. This proportion appears high at first, but is commensurate with the findings of Samsonidze et al.[22] who showed that semi-conducting nanotubes dissolved preferentially in ODA; if similar effects are present in Triton-dissolved nanotubes, we would anticipate the result described here. If the spherical model is used, the proportion of semi-conducting nanotubes is much higher (95 ± 0.5 %).

In order to validate these results, we compared the measurements to those obtained using optical absorption, a common technique for the determination of the proportions of metallic and semi-conducting SWNTs[23]. UV/Vis/NIR spectroscopy was performed using a Lambda 9 Perkin Elmer spectrometer by drop-casting the solution on a glass slide and heating to 200 ºC for 1 h. The spectrum obtained, following baseline correction to remove background absorption due to the π plasmon, is shown in figure 3. An average diameter of 1.2 nm was estimated from the first peak, corresponding to

the v_1-c_1 singularity transition of semi-conducting SWNTs[24]. The ratio between the first and the third peak (corresponding to the v_1-c_1 singularity transition of metallic SWNTs) indicated a population of 79.6% (± 3 %) of semi-conducting SWNTs in the Triton solution before dielectrophoresis, which is in good agreement with our calculations for nanotubes when considering the semi-conducting SWNTs as rod-like elements as opposed to spherical objects. This assumption is valid when we consider that the model of nanotubes as spheres only hold for ballistic transport, which does not take place in semi-conducting SWNTs. This result also suggests that it is possible to determine the proportion of metallic and semi-conducting SWNTs in a suspension using only two impedance measurements, at frequencies chosen so as to collect both types at low frequency, and only metallic nanotubes at high frequency. Such a method has significant potential to greatly simplify the methods available today for the analysis of SWNTs pre- and post-separation.

In conclusion, we have demonstrated a rapid, precise and low-cost method for performing dielectric spectroscopy on single-walled carbon nanotube suspensions. This has allowed the determination of the dielectric properties of the SWNTs with good precision, and also the rapid determination of the proportions of metallic and semi-conducting SWNTs. Moreover, the precision afforded by the system presented here has potential for providing more accurate control over defining the optimal conditions for the large-scale separation of SWNTs that is required for the adoption of nanotubes as a material of choice in the semi-conductor industry.

Acknowledgement. This work was supported by EPSRC grant number GR/S44594. The authors thank Ewa Borowiak-Palen for her help on optical absorption and Alan Dalton for supplying the SWNTs suspension.

Supporting Information Available. We used a suspension of SWNTs (laser ablation) in a TritonX-100 solution (0.6 % Triton X-100, 0.16 % SWNTs in weight). The suspension has been sonicated and then centrifuged for 2 h at 16000 g to remove the large impurities (catalyst particles, amorphous carbon). The electrodes were prepared from a 20 nm chromium film deposited by sputtering on glass. The patterning was made by photolithography and etching. The average gap between the electrodes is 10 μm.

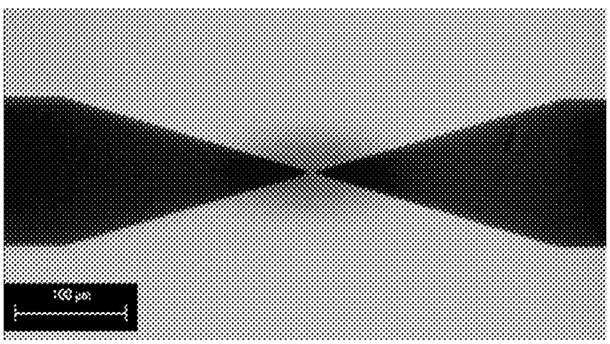
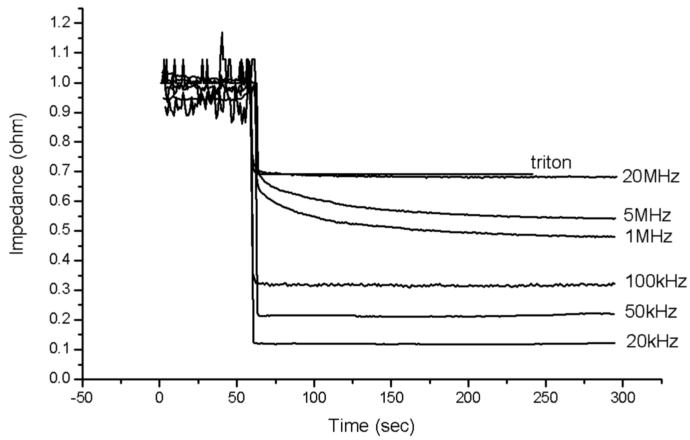


Figure 1. SWNTs collection on microelectrodes.

An optical microscope image showing SWNTs collected by dielectrophoresis after 3 min application of a 10 $V_{pk\text{-}pk}$, at 10 MHz. The collection of the carbon nanotubes can be observed aligned along the electric field lines in between the gap.

a)



b)

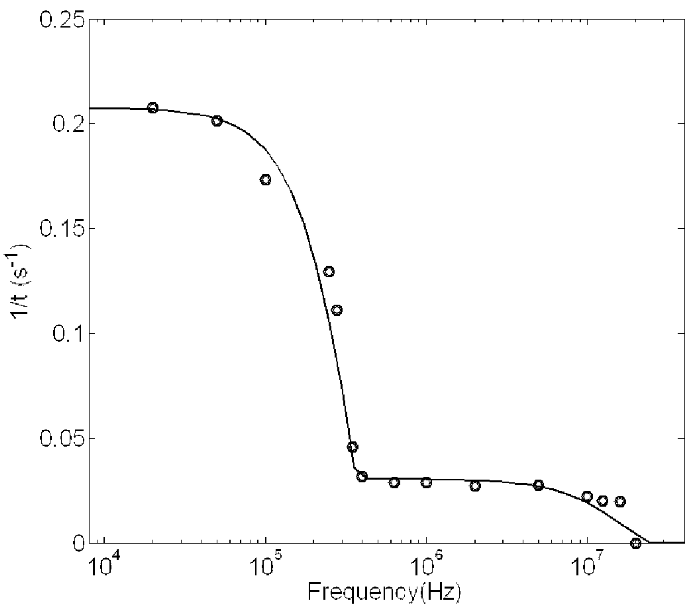


Figure 2. Experimental results of impedance measurements.

a) A plot of the impedance between the electrodes as a function of time for each frequency, after normalization against initial value. The time point at which the nanotubes suspension is applied to the electrodes is clearly visible at 60 s. **b)** A plot of reciprocal collection time constant as a function of frequency (dots) with the predicted value (solid line) using a model of the dielectrophoretic collection of two types of nanotubes superimposed.

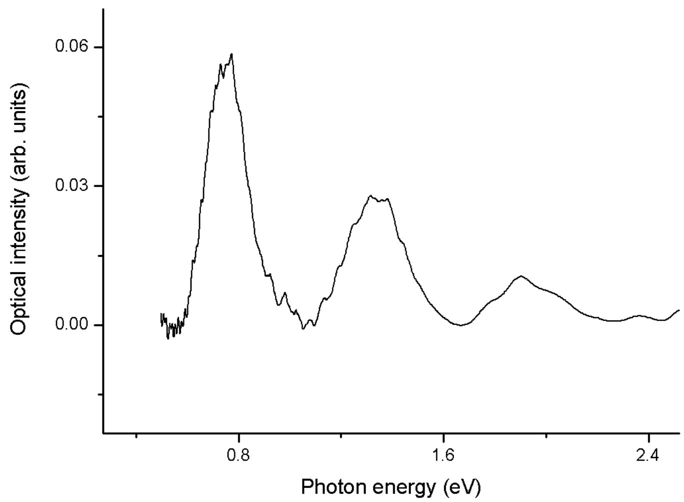


Figure 3. Optical absorption spectra of the Triton solution before dielectrophoresis.

Background absorption due to the π plasmon was corrected and abscissa scale converted from nanometer to commonly used electron-Volt.

References

1. Iijima, S. *Nature (London)* **1991**, 354, 56-57.

2. Rakov, E.G. *Russian Chemical Reviews* **2001,** 70, 10, 827-863.

3. Dresselhaus, M.S.; Dresselhaus, G.; Avouris, Ph. *Carbon nanotubes: synthesis, structure, properties and applications.* Eds.; Springer: New York, 2001.

4. Saito, R.; Dresselhaus, G.; Dresselhaus, M.S. *Physics properties of carbon nanotubes.* Imperial College, London, U.K., 1998.

5. Dekker, C. *Phys. Today* **1999** 52, 5, 22-28.

6. Pohl, H.A. *Dielectrophoresis.* Cambridge Univ. Press: Cambridge, England, 1978.

7. Hughes, M. *Nanoelectromechanics in engineering and biology.* London; Boca Raton, FL: CRC, 2003.

8. Krupke, R.; Hennrich, F.; von Löhneysen, H.; Kappes, M.M. *Science* **2003**, 301, 344-347.

9. Lee, D.S.; Kim, D.W.; Kim, H.S.; Lee, S.W.; Jhang, S.H.; Park, Y.M.; Campbell, E.E.B. *Appl. Phys. A* **2005**, 80, 5-8.

10. Krupke, R.; Hennrich, F.; Weber, H.B.; Kappes, M.M.; von Löhneysen, H. *Nano Lett.* **2003**, 3, 8, 1019-1023.

11. Strano, M.S.; Doorn, S.K.; Haroz, E.H.; Kittrell, C.; Hauge, R.H.; Smalley, R.E. *Nano Lett.* **2003**, 3, 8, 1091-1096.

12. Kataura, H.; Kumazawa, Y.; Maniwa, Y.; Umezu, I.; Suzuki, S.; Ohtsuka, Y.; Achiba, Y. *Synthetic Metals* **1999**, 103, 2555-2558.

13. Keszler, A.M.; Nemes, L.; Ahmad, S.R.; Fang, X. *Journal of optoelectronics and advanced materials* **2004**, 6, 4, 1269-1274.

14. Bachilo, S.M.; Strano, M.S.; Kittrel, C.; Hauge, R.H.; Smalley, R.E.; Weisman, R.B. *Science* **2002**, 298, 2361-2366.

15. Bakewell, D. J.; Morgan, H. Meas. *Sci. Technol.* **2004**, 15, 254-266.

16. Broche, L.; Labeed, F.H.; Hughes, M.P. *Phys. Med. Biol.* **2005**, 50, 2267-2274.

17. Morgan, H.; Green, N.G. *Journal of electrostatics* **1997**, 42, 279-293.

18. Benedict, L.X.; Louie, S.G.; Cohen, M.L. *Phys. Rev. B* **1995**, 52, 8541-8549.

19. Hughes, M.P.; Morgan, H.; Flynn, M.F. *J. Colloid Interf. Sci.* **1999**, 220, 454-457.

20. Hughes, M.P.; Green, N.G. *J. Colloid Interf. Sci.* **2002**, 250, 266-268.

21. Hughes, M.P. *J. Colloid Interf. Sci.* **2002**, 250, 291-294.

22. Samsonidze, Ge.G.; Chou, S.G.; Santos, A.P.; Brar, V.W.; Dresselhaus, G.; Dresselhaus, M.S.; Selbst, A.; Swan, A.K.; Ünlü, M.S.; Goldberg, B.B.; Chattopadhyay, D.; Kim, S.N.; Papadimitrakopoulos, F. *Appl. Phys. Letters* **2004,** 85, 1006-1008.

23. Kataura, H.; Kumazawa, Y.; Maniwa, Y.; Umezu, I.; Suzuki, S.; Ohtsuka, Y.; Achiba, Y. *Synthetic Metals* **1999,** 103, 2555-2558.

24. Liu, X.; Pichler, T.; Knupfer, M.; Golden, M.S.; Kink, J.; Kataura, H.; Achiba, Y. *Phys. Rev. B* **2002**, 66, 045411/1-045411/8.

X. RELATED PROCEEDINGS APPENDIX

Not Applicable